# DOCUMENT RESUME

| | |
|---|---|
| ED 422 676 | EC 306 626 |
| AUTHOR | Kratochwill, Thomas R.; Johnston, Hugh; Racine, Caroline N. |
| TITLE | The Development of the Obsessive-Compulsive Disorder School Impact Scale: An Investigation Using the Accuracy-Reliability Paradigm for Scale Validation. |
| INSTITUTION | Wisconsin Center for Education Research, Madison. |
| SPONS AGENCY | Office of Special Education and Rehabilitative Services (ED), Washington, DC. |
| PUB DATE | 1998-09-01 |
| NOTE | 205p. |
| CONTRACT | H-023B60024 |
| PUB TYPE | Reports - Descriptive (141) |
| EDRS PRICE | MF01/PC09 Plus Postage. |
| DESCRIPTORS | *Behavior Disorders; Behavior Patterns; Behavior Rating Scales; Elementary Secondary Education; *Emotional Disturbances; *Psychometrics; *Test Construction; *Test Reliability; *Test Validity |
| IDENTIFIERS | Expertise; *Obsessive Compulsive Behavior |

ABSTRACT

This report describes the development of a rating scale and direct observation measure to monitor the treatment of Obsessive Compulsive Disorder (OCD) in the home or school setting. The study applied an accuracy-reliability paradigm to validate the newly created rating scale and observation measure. An incontrovertible index using a scripted performance captured on videotape was developed. Participants included 10 student actors, an adult actor (the teacher), six trained graduate student observers, and nine experts in the field of OCD. Major findings included: (1) OCD expert ratings and observations ranged from moderately accurate to accurate; (2) graduate student observers were fairly reliable using the direct observation system and rating scale; (3) moderate to moderately high correlations were found between observations and ratings; (4) the direct observation system demonstrated sensitivity to change in the expected direction for the four target behaviors; and (5) the rating scale demonstrated sensitivity to change when the experts completed it, but not when the graduate students used it. Individual chapters of the report provide an introduction, a literature review, methodology explanation, presentation of results, and a discussion. Appendices include letters, coding information, the training manual, and the raw data. (Contains 59 references.) (DB)

*************************************************************************
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
***********************************************************************

ED 422 676

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

September 1, 1998

---

# The Development of the Obsessive-Compulsive Disorder School Impact Scale: An Investigation Using the Accuracy-Reliability Paradigm for Scale Validation

**Thomas R. Kratochwill & Hugh Johnston, Faculty Investigators**

**Caroline N. Racine, Student Investigator**

Final report to the U.S. Department of Education – Office of Special Education & Rehabilitative Services, Student-Initiated Research Grant No. H-023B60024

---

**Wisconsin Center for Education Research**
**School of Education University of Wisconsin–Madison**

EC 306626

## Table of Contents

Acknowledgments iv

List of Tables vi

List of Figures viii

List of Abbreviations ix

Abstract xi

Chapter 1: Introduction 1

Assessment of OCD 2

Focus and Conceptual Framework 5

Chapter 2: Literature Review 8

Overview of OCD 8

Clinical Course of OCD 9

Prevalence of OCD 10

Complications from OCD 11

Associated (Comorbid) Disorders 13

Etiology of OCD 14

Assessment of OCD 17

Rating Scale Assessment 20

Direct Observation Assessment 28

Reliability and Accuracy of Ratings and
    Direct Observation Measures 29
Relationship between Ratings and Direct Observations 33
Summary and Predictions 38
Chapter 3: Method 40
    Participants 40
    Materials 40
    Procedure 43
    Data Collection 45
    Research Design and Data Analysis 47
Chapter 4: Results 49
    Prediction 1 51
    Prediction 2 68
    Prediction 3 73
    Prediction 4 76
    Prediction 5 81
    Prediction 6 85
Chapter 5: Discussion 88

Generalizations and Implications of the Findings 89

Reliability of Ratings and Direct Observations 89

Accuracy of Ratings and Direct Observations 93

Relationship between Direct Observations and Ratings 99

Sensitivity to Change 101

Limitations and Suggestions for Future Research 103

Conclusion 106

References 110

Appendices 120

Appendix A: Consent Letters 120

Appendix B: Behavior Codes, OCD-RS, OCD-DOS 134

Appendix C: Training Manual 144

Appendix D: Raw Data 188

# List of Tables

Table 1. Research studies investigating psychometric properties of the LOI-CV and CY-BOCS

Table 2. Data Collection Procedure After Observation Training

Table 3. Descriptive Statistics by Target Behavior for OCD Expert Observations Using the OCD-DOS on Videotapes 1 and 2

Table 4. Descriptive Statistics by Target Behavior for OCD Expert Ratings Using the OCD-RS on Videotapes 1 and 2

Table 5. Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #1

Table 6. Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #2

Table 7. Descriptive Statistics by Target Behavior across Graduate Student Raters Using the OCD-RS on Videotapes 1 and 2

Table 8. Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #1

Table 9. Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #2

Table 10. Descriptive Statistics by Target Behavior across Graduate Student Observers Using the OCD-DOS on Videotapes 1 and 2

Table 11. Pearson Correlations between the Performance Standard's and the Graduate Student's Observations on the OCD-DOS on Videotapes 1 and 2

Table 12. Pearson Correlations between the Performance Standard's and the Graduate Student's Ratings on the OCD-RS on Videotapes 1 and 2

Table 13. Pearson Correlations between Graduate Student's OCD-DOS and His/Her Own OCD-RS Rating on Videotapes 1 and 2

Table 14. Pearson Correlations between and OCD Expert's OCD-DOS and His/Her Own OCD-RS Rating on Videotapes 1 and 2

Table 15. Effect Sizes for OCD Experts and Graduate Students by Target Behavior on the OCD-DOS and the OCD-RS

# List of Figures

Figure 1. Conceptual matrix for the development of assessment Instruments by purpose

Figure 2. OCD experts' and graduate students' mean OCD-DOS observations compared to the performance standard's OCD-DOS observations for videotapes #1 and #2

Figure 3. OCD experts' and graduate students' mean OCD-RS ratings compared to the performance standard's ratings for videotapes #1 and #2

Figure 4. Individual OCD expert's observations on the OCD-DOS compared to the performance standard for videotape #1 and #2

Figure 5. Individual graduate student's observations on the OCD-DOS compared to the performance standard for videotape #1 and #2

Figure 6. Individual OCD expert's OCD-RS ratings compared to the performance standard for videotape #1 and #2

Figure 7. Individual graduate student's OCD-RS ratings compared to the performance standard for videotape #1 and #2

Figure 8. Developmental phases for future research

## List of Abbreviations

| | |
|---|---|
| CBCL-DOF | Child Behavior Checklist-Direct Observation Form |
| CMI | Clomipramine |
| CPRS-OC | Comprehensive Psychopathological Rating Scale-Obsessive Compulsive subscale |
| CY-BOCS | Children's Yale-Brown Obsessive Compulsive Scale |
| DSM-IV | Diagnostic and Statistical Manual of Mental Disorder-Fourth Edition |
| ECA | Epidemiological Catchment Area |
| E/RP | Exposure and Response Prevention |
| ES | Effect Size |
| LOI | Leyton Obsessional Inventory |
| LOI-CV | Leyton Obsessional Inventory-Child's Version |
| NIMH-G | National Institute of Mental Health-Global scale |
| NIMH-OC | National Institute of Mental Health-Obsessive Compulsive scale |
| OCD | Obsessive-Compulsive Disorder |
| OCR | Obsessive-Compulsive Rating scale |

| | |
|---|---|
| OCD-RS | Obsessive-Compulsive Disorder-Rating Scale |
| OCD-DOS | Obsessive-Compulsive Disorder-Direct Observation System |
| PII | Problem Identification Interview |
| PS | Performance Standard |
| RMS | Root Mean Square |
| SRI | Serotonin Reuptake Inhibitor |
| SMAS | Selective Mutism Analysis Scale |
| SMOP | Selective Mutism Observation Protocol |
| SSBS | School Social Behavior Scale |
| SSRS-T | Social Skills Rating Scale-Teacher form |
| SSRI | Selective Serotonin Reuptake Inhibitor |
| Y-BOCS | Yale-Brown Obsessive Compulsive Scale |

## Abstract

Once thought to be rare, Obsessive-Compulsive Disorder (OCD) afflicts more than 2% of the population. This percentage may be an underestimation given the secrecy of this disorder. In fact, approximately one-third to one-half of all adults with OCD report that their OCD developed in childhood. Thus, the knowledge, identification, and ability to monitor interventions is crucial in both the home and school settings. Over the past decade, the investigation of successful treatment options has grown, and has provided many children and adolescents with some symptom alleviation. Unfortunately, the assessment methodology of OCD significantly lags behind that of current intervention options. Therefore, the assessment domain needs to be addressed. The purpose of this study was to develop a rating scale and direct observation measure that could be used in the home or school setting to monitor the treatment of OCD on a case-specific basis. Participants in this study were: (1) student actors (N=10); (2) an adult actor (i.e., the teacher) (N=1); (3) trained graduate student observers (N=6); and (4) experts in the field of OCD (N=9). The accuracy-reliability paradigm was used to validate the newly created rating scale and observation measure. To accomplish this task, an incontrovertible index was developed using a scripted performance captured on videotape. The research design and data analysis was descriptive and correlational. Several findings of this study were noteworthy. First, OCD expert ratings and observations ranged from moderately accurate to accurate. Second, graduate student observers were fairly reliable on the OCD-DOS and OCD-RS. Also, moderate to moderately high correlations were found between observations and ratings. Finally, the OCD-DOS demonstrated sensitivity to change in the expected direction for the four target behaviors. The OCD-RS demonstrated sensitivity to change when the experts completed it, but not when the graduate students used it. A detailed examination of these findings and their implications for researchers and practitioners is discussed.

# Chapter 1

## Introduction

Obsessive-Compulsive Disorder (OCD), originally thought to be a rare disorder, afflicts more than 2% of the population (Whitaker et al., 1990). Among adults who suffer from OCD, approximately one-third to one-half developed OCD when they were children or adolescents (March, Leonard, & Swedo, 1995). Realistically, OCD probably affects more individuals than current epidemiological studies report. Due to the hidden nature of OCD, it is reasonable to assume that it is an under-reported and under-identified disorder that, for a large majority of people, has origins in childhood. In many cases, early detection of this disorder has played a vital role in children's recovery from OCD (Querioz, Motta, Madi, Sossai, & Boren, 1981). Therefore, knowledge and identification of OCD by school psychologists is important to the alleviation of this disorder (Adams, Waas, March, & Smith, 1994).

Briefly, obsessions can be described as "persistent ideas, thoughts, impulses, or images that are experienced as intrusive and inappropriate and that cause marked anxiety or distress" (American Psychiatric Association, 1994, p. 418). Some typical childhood obsessions are fear of contamination, fear of something happening to themselves or loved ones (i.e., fire, death, illness), and need for exactness (Swedo, Leonard, &

Rapoport, 1990). Compulsions can be defined as "repetitive behaviors (e.g., hand washing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) the goal of which is to prevent or reduce anxiety or distress, not to provide pleasure or gratification" (American Psychiatric Association, 1994, p. 418). During childhood, it is not uncommon for children to demonstrate typical age-dependent ritualized behaviors that resemble OCD (i.e., ritualized play or bedtime rituals). However, these rituals are typically linked to developmental tasks and usually vanish by middle childhood (Evans, 1997; Ilg, Ames, & Baker, 1981; Leonard, Goldberger, & Rapoport, 1990). The obsessions and/or compulsions of children with OCD differ from developmentally-appropriate ritualized behaviors in respect to timing, content, and severity. That is, the obsessions and compulsions of children with OCD usually appear after the developmentally-appropriate time, seem odd, and produce an inability to function (March, Leonard, & Swedo, 1995).

## Assessment of OCD

A review of the childhood-onset OCD assessment literature revealed several shortcomings in the currently used assessment devices. These weaknesses are: (a) the children's scales have not been designed specifically for children in that they have been downward extensions of the adult scales; (b) the scales have untested reliability and validity; (c) they are primarily clinician-administered interviews or rating scales; and (d) the instruments have a lack of symptom specificity (Como, 1997). Additionally, direct behavioral observations for assessment and intervention monitoring have not been used on a consistent or frequent basis (Henin & Kendall, 1997; Taylor, 1995).

The most glaring shortcoming of all current OCD assessment devices has been the failure to use appropriate validity indices. This lack of appropriate validity estimates for OCD assessment devices significantly hinders the utility of these instruments. All of the OCD measures that report validity estimates have used a very traditional method for establishing content, construct, and criterion-related validity. In essence, investigators have tried to validate OCD assessment instruments by using a traditional scale validation paradigm that seeks to assess the existence of a hypothetical construct, or latent trait. While this traditional validation paradigm represents one way to establish the validity of instruments, perhaps an alternative scale validation model may be more useful in establishing the validity of OCD assessment measures. In this context, an effort should be made to validate the OCD assessment devices as measures of behavior instead of measuring a latent trait, or state of being. In other words, OCD can be considered a behavioral response class consisting of numerous individual, molecular behaviors (i.e., obsessions and compulsions). The purpose of validating

OCD assessment measures as behavioral assessment measures is to ensure that the resulting scores describe actual, observable behavior (i.e., motor responses, cognitive and affective events, and physiological responses).

Foster and Cone (1995) assert that two types of representational validity, content validity and accuracy, be established for behavioral measures. Content validity assesses the extent to which the instrument measures what it purports to measure. Specifically, content validity examines the extent to which a measure operationally defines the behavior/response class that is being assessed. Also, scores obtained from the measure should depict the actual behavior. Accuracy is a measure of the instrument's ability to reflect "true" behavior. Establishing an instrument's accuracy is done by comparing scores on a given instrument to those of an incontrovertible index. An incontrovertible index is an illustration of the target behavior(s) that the investigator deems as "truth." This incontrovertible index can be generated by mechanically generated responses, naturalistic observations, and/or controlled stimuli (i.e., videotaped naturalistic behavior or scripted performance). Once accuracy of the instrument is established, convergent and discriminant validity are no longer necessary (Foster & Cone, 1995).

A review of psychopharmacological treatment studies and psychosocial interventions for children who have OCD underscores the inadequacies of the current measures to reliably and accurately assess behavior change. In terms of both medication and psychosocial intervention, clinicians and researchers assert that there has been a positive change; however, they often fail to provide consumers with the data on what behaviors have changed.

Over the past decade, the investigation of successful treatment options for children and adolescents with OCD has multiplied. Fortunately, exposure and response prevention (E/RP) behavior therapy, drug therapy, or a combination of these therapies have afforded many children and adolescents some overall relief; and for a few, full recovery (Leonard, Swedo, March, & Rapoport, 1995). Unfortunately, the assessment methodology of OCD significantly trails that of current treatment options.

## Focus and Conceptual Framework

The focus of this dissertation was twofold. First, two outcome assessment measures (i.e., a rating scale and direct observation measure) to monitor psychopharmacological and/or behavioral interventions were created. Second, the reliability and accuracy of the newly-created rating scale and direct observation measure was established through an alternative scale validation model called the accuracy-reliability paradigm (Racine, 1994).

Given that the psychometric properties of the various childhood-onset OCD scales are suspect, there is a need to develop a scale that is sensitive to the topography of a certain behavior. Accuracy, a concept closely related to reliability and validity, has rarely been addressed in the evaluation of rating scales or direct observation measures. Cone (1981) indicated that the establishment of an instrument's accuracy is the most important feature. Unfortunately, researchers have often assumed accuracy from inter-rater reliability or test-retest reliability indices. However, a measure that is reliable is not necessarily accurate, but a measure that is accurate is inevitably reliable. Cone (1992) asserts, "To establish an instrument's accuracy, whether for occurrence, cross-setting, or other, it is necessary to have two things: (a) a set of rules/procedures, preferably written, for using the instrument, and (b) an incontrovertible index against which to compare data produced by the instrument." (p. 24). Thus, an incontrovertible index was be established for purposes of this research by using two videotaped vignettes of scripted performances from an adolescent who exhibited four frequently observed compulsions. Two UW-Madison experts came to consensus on what they observed while watching the two videotapes. Hence, the performance standard (i.e., incontrovertible index) was derived. Then, the trained experts and experts in the field of OCD viewed these videotapes and completed the newly created rating scale and direct observation measure. If the participants converged on what they observed, then the scale is accurate. In addition, written, explicit instructions were developed in accordance with Cone's statement via a training manual. Specific instructions were an important part because they are almost always never used with existing instruments (Taylor, 1995).

For purposes of this study, the following literature was reviewed: (1) general OCD issues; (2) assessment issues related to childhood-onset OCD; (3) the reliability of rating scales and direct observation measures; and (4) the accuracy of behavioral measures. This relevant literature is reviewed in Chapter 2.

# Chapter 2

## Literature Review

### Overview of OCD

According to the Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV) (1994) an individual must have five diagnostic characteristics to meet the criteria for clinical OCD. These characteristics, as outlined by the DSM-IV (1994), are as follows:

> A. Either obsessions or compulsions:
>
> *Obsessions* as defined by (1), (2), (3), and (4):
> (1) recurrent and persistent thoughts, impulses, or images that are experienced, at some time during the disturbance, as intrusive and inappropriate and that cause marked anxiety or distress
>
> (2) the thoughts, impulses, or images are not simply excessive worries about real-life problems
>
> (3) the person attempts to ignore or suppress such thoughts, impulses, or images, or to neutralize them with some other thought or action
>
> (4) the person recognizes that the obsessional thoughts, impulses, or images are a product of his or her own mind (not imposed from without as in thought insertion)
>
> *Compulsions* as defined by (1) and (2):
> (1) repetitive behaviors (e.g., hand washing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) that the person feels driven to perform in response to an obsession, or according to rules that must be applied rigidly
>
> (2) the behaviors or mental acts either are aimed at preventing or reducing distress or preventing some dreaded event or situation;

> however, these behaviors or mental acts either are not connected in a realistic way with what they are designed to neutralize or prevent or are clearly excessive
>
> B. At some point during the course of the disorder, the person has recognized that the obsessions or compulsions are excessive or unreasonable. **Note**: This does not apply to children.
>
> C. The obsessions or compulsions cause marked distress, are time consuming (take more than 1 hour a day), or significantly interfere with the person's normal routine, occupational (or academic) functioning, or usual social activities or relationships.
>
> D. If another Axis I disorder is present, the content of the obsessions or compulsions is not restricted to it
>
> E. The disturbance is not due to the direct physiological effects of a substance (e.g., a drug of abuse, medication) or a general medical condition. (p. 422-423).

OCD is diagnosed along Axis I and is considered to be an anxiety disorder. Childhood-onset OCD must meet the same criteria as specified for adults in the DSM-IV; however, children may or may not have insight into their obsessions and/or compulsions.

## Clinical Course of OCD

The topography of childhood-onset OCD varies on an individual basis as well as intra-individually across time. In other words, the course of childhood-onset OCD does not seem to follow any typical developmental trajectory nor are certain obsessions and/or compulsions predictive of a certain developmental pathway (March & Leonard, 1996). In a prospective follow-up study by Rettew, Swedo, Leonard, Lenane and Rapoport (1992), qualitative information regarding OCD symptoms across time was reported for a sample of children and adolescents (N=79). Many interesting findings were reported: (a) no significant age-related trends were discovered in terms of symptom type; (b), washing and checking compulsions were exhibited by 47% of the sample at some point during the course of their disorder; (c) symptom constellation changed over time and varied on content and severity dimensions; and (d) children 6-years old and younger were more likely to have compulsions. These authors also noted that pure obsessive OCD was observed less frequently than pure compulsive OCD. The age of onset data from this sample yielded an early onset type and an adolescent onset type. Those children who were diagnosed with early onset OCD were twice as likely as those with adolescent onest OCD to be male and to have had a positive family history of OCD or Tourette's Syndrome.

## Prevalence of OCD

Prevalence estimates of OCD usually vary between 2% and 3% of the general population. For example, the Epidemiology Catchment Area (ECA) study reported a 2.5% of lifetime prevalence in the population at large (Robins et al., 1984). Flament and colleagues (1988) conducted a survey of 5,596 high school students, and 20 adolescents were identified as having OCD (0.4%). When the sample was weighted, the estimated prevalence rate for adolescents with OCD was 1%. In addition, the majority of childhood and adolescent OCD studies indicate that most children diagnosed with OCD show symptoms during 10 to 14 years of age and another one-third report symptoms before 9 years of age (Riddle et al., 1990; Swedo et al., 1989). Flament et al. (1988) reported a six month prevalence rate of approximately 1 in 200 children and adolescents. The true prevalence of childhood-onset OCD is difficult to ascertain given that there have been no appropriate epidemiological child and adolescent studies (Carter, Pauls, & Leckman, 1995).

## Complications from OCD

Complications from OCD can affect children and adolescents both at school and home.

Impact of OCD at school. Although the prevalence rate of OCD in children and adolescents is not extraordinary, it is believed to be 20 to 40 times more common than previously reported (Clarizio, 1991). The secretive nature of this problem makes it difficult to identify those who have OCD. Although there has been no systematic investigation as to how youth with OCD function in school, it is important for school personnel to be knowledgeable about OCD and be able to critically evaluate referrals for special education. Due to the individual variability of OCD across students, OCD will impact students in different ways. Among the most common school problems (from time of onset) that students with OCD experience are: (a) poor peer relations; (b) high absentee rates; (c) a decline in scholastic aptitude and overall school functioning; (d) an increase in drug and alcohol abuse, and (e) lack of attention/distractibility (Adams et al., 1994; Clarizio, 1991; Johnston & Fruehling, 1997). Many of these students also are referred and placed in special education due to their inability to function in a regular education classroom; however, other students with mild to moderate OCD remain in the regular education setting.

Impact of OCD at home. Each family who has a child or adolescent with OCD copes differently. VanNoppen, Pato, and Rasmussen (1993) delineate five different coping strategies typically adopted by families who have a family member with OCD. First, there are those families who help with the child's rituals in an effort to maintain the status quo. Second, there are those families who do not assist in the rituals, but allow these compulsions to occur. Third, there are those families who ignore or fail to acknowledge the obsessions and compulsions. Fourth, many families respond differentially to the family member with OCD. Finally, there are those families who respond inconsistently to the family member with OCD. Children and adolescents with OCD often try hard to keep their OCD behaviors hidden from other family members. Often times, children and adolescents feel isolated and ashamed because they realize that their obsessions and/or compulsions are irrational. Children and adolescents with OCD typically are dependent on their parents for many of their needs. This dependence often is troublesome for adolescents because OCD hinders the separation and individuation process that most adolescents experience during the teenage years. The key to coping with a child or adolescent who has OCD is to seek help, allow for open family discussion about OCD, and provide consistent discipline and routines (VanNoppen et al., 1993).

Associated (Comorbid) Disorders

The comorbidity of OCD with other disorders is very common. In fact, child and adolescent studies that report on comorbidity of their samples reveal similar data in regard to OCD and associated disorders (i.e., Swedo et al., 1989; Swedo, Leonard, & Rapoport, 1992; and Riddle et al., 1990). Data from Swedo and colleagues (1989,1992) reveal that only 25% of the sample ($\underline{N}$ = 140) had OCD as their sole diagnosis. The data ($\underline{N}$=70) from Swedo et al. (1989) provide an excellent illustration of psychiatric disorders commonly associated with OCD: tic disorder (30%), major depression (26%), specific developmental disorder (24%), simple phobia (17%), overanxious disorder (16%), adjustment disorder with depressed mood (13%), oppositional defiant disorder (11 %), attention deficit disorder (10%), conduct disorder (7%), and separation anxiety disorder (7%). These results highlight the need to consider comorbid conditions when diagnosing children and adolescents with OCD as dual diagnoses will affect intervention planning.

## Etiology of OCD

Currently, there are many different hypotheses as to the cause of OCD. The only definitive statement that can be offered with any confidence is that the cause of OCD remains nebulous. Nonetheless, there are various models and hypotheses of OCD which help clinicians to pursue various treatment options. Several contemporary models of OCD are presented briefly to illustrate the scope and diversity of current explanations of OCD. Most of these models have focused primarily on adults; consequently, researchers are left to infer that these adult models also apply to children and adolescents.

Neurobiological models. Although there are several neurobiological and neuropsychological models of OCD currently being investigated, the serotonin connection has received the most attention. Evidence that serotonin, a neurotransmitter, is involved in OCD was first obtained when researchers discovered that clomipramine (CMI), a tricyclic antidepressant as well as a serotonin (and other monoamine) reuptake inhibitor (SRI), reduced OCD symptoms in adults. This finding has been replicated with children and adolescents ( i.e., Flament et al., 1985; Leonard et al., 1989; and DeVeaugh-Geiss et al., 1992). Additional evidence that implicates serotonin dysfunction in the etiology of OCD has been obtained in: (a) direct assays of serotonin and its metabolites; (b) drug challenge studies; and (c) the efficacy of selective serotonin reuptake inhibitors (SSRIs) (Johnston & March, 1992). Research also has demonstrated that individuals with OCD have reduced levels of serotonin in the synaptic cleft as well as post-synaptic hyperactivity. Thus, medications that increase serotonin in the synapse by blocking its reuptake into the presynaptic neuron have alleviated some OCD symptoms (Johnston & March, 1992; Piacentini et al., 1992). This model continues to remain viable; however, it does not point to a sole cause of OCD because some individuals with OCD do not improve when taking CMI or SSRIs.

Recently, researchers have been investigating the possibility that some forms of childhood-onset OCD may be caused by a viral infection. In reviewing the literature and clinical cases, these investigators discovered that two general patterns of OCD emerged among children and adolescents (Allen et al., 1995; Leonard, 1995; Swedo, 1994). Some children had an insidious onset of OCD that waxed and waned over time. Other children experienced a sudden onset of OCD that remitted between OCD episodes. For this latter group of children, there appeared to be an association between OCD onset and Group A Beta Hemolytic Streptococcal infection, or "strep throat" (Allen et al., 1995; Leonard, 1995; Swedo, 1994). These researchers have likened this process to that of Sydenham's chorea, a form of rheumatic fever, in that, " some patients are genetically predisposed to marshal an autoimmune response to Group A Beta Hemolytic Streptoccus when they form autoantibodies that are misdirected at various parts of their bodies, such as the heart (which can result in rheumatic carditis) or the brain (which can give rise to chorea)" (Leonard, 1995, p. 13). Hence, Sydenham's chorea might serve as a medical model for OCD. Although this model is promising, more research is needed.

Given the genetic transmission of possible serotonin dysfunction and autoimmune responses, it is not surprising that genetic factors are hypothesized to play a role in the development of OCD as well. Family and twin studies have been conducted to provide evidence of OCD heritability. Rettew et al. (1989) (*N* = 70) reported that 25% of those children and adolescents with OCD had a positive family history of OCD. Most of the family and twin studies report that OCD is familial; however, these studies are not without their methodological weaknesses (Rasmussen, 1994).

*Behavioral model*. Some believe OCD is a learned behavior that has resulted from classical conditioning and operant learning. This theory is referred to as a two-factor conditioning model (Johnston & March, 1992). For example, a new obsession (i.e., unconditioned stimulus) occurs, the individual experiences anxiety, and then the individual performs a behavior which the individual thinks will reduce this anxiety (i.e., compulsion). This unconditioned stimulus becomes a conditioned stimulus, and operant learning results. Environmental events and cues that surround the conditioned behavior often prompt other behaviors to arise, a process commonly referred to as stimulus generalization. Given that behavior therapy has been successful in symptom alleviation for some individuals, this behavioral model also remains a reasonable explanation as to the cause of OCD. Although, this theory fails to explain why some children develop OCD and others do not.

## Assessment of OCD

A best practice approach to the clinical diagnosis of OCD involves the collection of information using a variety of methods and informants across settings (Achenbach, McConaughy, & Howell, 1987; Francis & Gragg, 1996; Hnnin & Kendall, 1997). In the assessment of any anxiety disorder, instruments should: (a) be reliable and valid across multiple symptom domains; (b) differentiate symptom clusters; (c) assess the severity of the problem; (d) make use of multiple informants; and (e) be sensitive to behavior change (Stallings & March, 1995). Stallings and March (1995) also assert that assessment devices should facilitate communication among professionals especially in light of the current trend toward a multi-disciplinary team approach to assessment. Assessment purpose also should be delineated and considered prior to the actual assessment. For example, a clinician should reflect on whether the purpose of assessment is for diagnosis, treatment planning, and/or treatment monitoring (Foster & Cone, 1995; Kratochwill & McGivern, 1997). The delineation of assessment purpose will help guide the clinician in determining what type of assessment instrument to use (see Figure 1). Although, OCD has a controversy-free nosology, OCD presents itself with considerable inter-individual heterogeneity (March, Johnston, & Greist, 1990; Rettew et al., 1992). Thus, practitioners using various assessments must make a concerted effort to account for symptom variability within OCD (Stallings & March, 1995). Clinicians also need to consider comorbid conditions as this will impact choice of treatment. The methods most commonly used by clinicians to assess OCD include the following: a diagnostic interview (structured, semi-structured, and/or unstructured), rating scales (self and/or other), self-monitoring (e.g., behavioral diary method), and direct behavioral observations (Henin & Kendall, 1997; Taylor, 1995; Wolff & Wolff, 1991). Due to time efficiency and cost effectiveness, the most relied upon method in clinical practice is most likely the completion of rating scales. Even though several assessment instruments have been developed over the course of the past decade, " no one scale has emerged as the most broadly accepted scale for OCD " (Como, 1997, p.256). Additionally, researchers have been calling for the

| | RATING SCALE | DIRECT OBSERVATION |
|---|---|---|
| **SYMPTOM CONSTELLATION**<br>**Purpose**<br>**1. Treatment outcome monitoring**<br>**2. Diagnosis** | A rating scale format where OCD observable behavior can be ranked on certain dimensions<br>Example<br>OCD-Rating Scale (Appendix B)<br>CY-BOCS | Event and or duration recording completed by trained observers. Operational Definitions should be Parallel to Rating Scale Behaviors.<br>Example<br>OCD Behavior / Occurrence / Duration<br>1. tracing & retracing letters / 1 / 10 min<br>2 / 5 min<br>8 min |
| **FUNCTION**<br>**Purpose**<br>**1. Treatment Design** | Rate or Estimate antecedents, consequents, and sequential events when OCD behaviors occur.<br>Example<br>*Child Report*<br>1 How often do you trace letters or words over and over because of avoiding your work?<br>*Teacher Report*<br>1 Does the student seem to trace and retrace words or letters (or numbers) in order to avoid work completion? | Observe antecedents, consequents, and sequential events when OCD behaviors occur.<br>Example<br>A — B — C<br>target behavior ex. tracing letters over and over<br>A: group lesson<br>B: tracing tapping foot<br>C: avoids turning in work or answering Qs |

Figure 1. Conceptual Matrix for the Development of Assessment Instruments by Purpose

development of accurate OCD assessment devices and instruments that can provide data on a case specific basis (Como, 1997; Francis & Gragg, 1996). Given the focus of this dissertation, rating scale and direct observation assessment was reviewed.

## Rating Scale Assessment

Given the widespread use of rating scales among clinicians, it is important to be cognizant of the technical adequacies of the most commonly used rating scales. The two most commonly used rating scales, the Leyton Obsessional Inventory-Children's Version (LOI-CV) and the Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) are reviewed in depth. Unfortunately, content validity and accuracy of these scales have not been documented. Consequently, traditional estimates of validity are reported. Both of these scales are purported to have multiple roles in the assessment of OCD; however, they are most commonly used for diagnosis, treatment design, and treatment monitoring.

### The Leyton Obsessional Inventory-Child Version (LOI-CV).

The LOI-CV (Berg, Rapoport, & Flament, 1986) is a downward extension of the adult LOI (Cooper & McNeil, 1968) which was designed to assess obsessionality in "houseproud" homemakers. Cooper (1970) adapted it for the assessment of OCD. The LOI-CV, like the LOI, is a 44-item, self-report measure that allows assessment of the following symptom domains: thoughts, checking, dirt and contamination, dangerous objects, cleanliness and tidiness, school work, order and routine, overconscientious, indecision, hoarding, meanness, and magic games. The LOI-CV involves a card-sorting task where the child is presented a symptom on a card, and then he/she drops the symptom card in a "yes" or a "no" box. Positive responses are re-administered to assess the dimensions of resistance (i.e., how hard it is to stop - 5-point Likert-type response format) and interference (i.e., how much symptom interferes with daily functioning- 4- point Likert-type response format). Total symptom scores are calculated by adding the number of positive symptoms cards; and resistance and interference scores are added according to their point values. There also is a 20-item LOI-CV using the same format that has been developed for epidemiological studies.

In terms of reported technical adequacies for the LOI-CV (44 item), data have been scant. In fact, a literature search conducted by the Obsessive-Compulsive Information Center at the Dean Foundation located in Madison, Wisconsin revealed only 19 articles that referenced the LOI-CV. Of these 19 articles, only one addressed the technical adequacies of this instrument. Berg et al. (1986) specifically tested the reliability and validity of the LOI-CV by studying three groups: (1) 26 adolescents (age range = 10-18 yrs., M= 14.3 yrs., 17 males, 9 females) who were diagnosed with OCD, (2) 28 adolescent controls (age range = 11-18 yrs., M = 13.7 yrs., 22 males, 6 females), who were matched for age, and (3) 14 psychiatric adolescents (age range = 10-17 yrs., M = 13.2; 10 males, 4 females) who had some obsessive symptoms, but OCD was not their primary diagnosis. Of the first group, 19 participated in a double-blind crossover drug treatment trial. A one-way ANOVA found significant group differences resulted for the number of "yes," resistance, and interference scores. The test-retest reliability of 10 patients across a placebo phase of 5 weeks demonstrated intraclass coefficients of .96, .97, and .94 for the "yes," resistance, and interference scores, respectively. Additionally, Pearson product-moment correlations were computed between the LOI-CV and other OCD rating scales. Correlations ranged from .77 to .89 for both the LOI-CV and the Obsessive Compulsive Rating (OCR, Rapoport et al., 1980), the Comprehensive Psychopathological Rating Scale-Obsessive Compulsive Subscale (CPRS-OC, Thoren et al., 1980), and the NIMH Obsessive Compulsive scale (NIMH-OC, Rapoport et al., unpublished). Correlations between the LOI-CV and the NIMH Global subscale (NIMH-G,Murphy et al., 1982) ranged from .69 to .77. Paired t-tests were also computed between the end of baseline and the end of treatment scores. Significant differences were reported, indicating that the LOI-CV can provide useful information about clinical change.

Three studies from the same literature search were identified for providing information on the LOI-CV (20 items) epidemiological version

Flament et al. (1988) gathered epidemiological information from 356 high school students. These researchers reported the sensitivity of the LOI-CV (20 items) as 75%, the specificity as 84%, and the predictive validity as 18%. The sensitivity percentage indicates the number of cases that test positive for the target disorder, and the specificity percentage means the number of cases that test negative for the target disorder. Verhulst and Koot (1992) assert that these measures of validity do not demonstrate the intrinsic properties of an instrument. In fact, these percentages will vary between samples. The predictive value indicates the likelihood that a positive test result will predict a certain disorder (Verhulst & Koot, 1992). Given the 18% predictive validity, Flament and researchers (1988) assert that the LOI-CV (20 items) is an adequate screening measure. Berg and colleagues (1988) reported an internal consistency estimate of .81 based on Cronbach's $\alpha$ coefficient. King, Myerson, Inglis, Jenkins, and Ollendick (1995) examined the reliability of the LOI-CV (20 items) in an Australian sample (N=1602) with an age range of 8 to 16 years. These researchers reported a Cronbach's $\alpha$ of .76. Two-week test-retest reliability was calculated using Total Obsessive scores and Total Interference scores for three different age cohorts (8-10 yr.-olds, 11-13 yr.-olds, and 14-16 yr.- olds) on 106 randomly selected participants. For the youngest cohort, .51 and .65 on Total Obsessive and Total Interference scores were calculated, respectively. For the middle cohort, 75 and .81, for Total Obsessive and Total Interference scores were derived, respectively. For the oldest cohort, .83 and .57 for Total Obsessive and Total Interference scores were found, respectively. Total Obsessive scores were observed to increase with age; however, this was not true for the Interference scores. (see Table 1 for Technical Information on LOI-CV.)

This dearth of tested reliability and validity for the LOI-CV (44-item and 20-item scales) makes it extremely hard to judge its technical adequacy as well as usefulness in the assessment process. Additionally, the studies that were found are difficult to compare given the lack of common statistics used to judge the instrument. Finally, it should be underscored that traditional validity studies are distinctly different from epidemiological validity studies, and caution should be rendered when trying to compare statistics across studies.

Given these caveats about the lack of reported technical adequacies, the advantages to the LOI-CV are: (a) it is the only self-report measure available for children; (b) it provides a measure of presenting symptoms, resistance, and interference; and (c) the card sort method affords clinical utility through observations. The disadvantages are: (a) it does not distinguish between obsessions and compulsions; (b) it is time consuming to administer; and (c) there is potential for an overabundance of false positives and false negatives (Berg, 1989).

Table 1

Research studies investigating psychometric properties of the LOI-CV and CY-BOCS

| RESEARCH STUDY | RELIABILITY | | | VALIDITY | | |
|---|---|---|---|---|---|---|
| | Internal Consistency | Test-Retest | Inter-Rater | Content | Criterion | Construct |
| **LOI-CV (44)** | | | | | | |
| Berg et al. 1986<br>(N = 54) | NR | .96 Y<br>.97 R<br>.94 I<br>(n=10) | NR | NR | .77-.89 OCR<br>CPRS-OC<br>NIMH-OC<br>.69-.77 NIMH-G | NR |
| **LOI-CV (20)** | | | | | | |
| Berg et al., 1988<br>(N= 5596) | .81 | NR | NA | NR | NR | NR |
| Flament et al., 1988<br>(N = 356) | NR | NR | NA | NR | 75% Sensitivity<br>84% Specificity<br>18% Predictive | NR |
| King et al., 1988<br>(N= 1602) | .76 | .71<br>(n=106) | NA | NR | NR | NR |
| | | | | | | |
| **CY-BOCS** | | | | | | |
| Hanna 1995<br>(N= 31) | NR | NR | NR | NR | .93 NIMH-G | NR |
| Scahill et al (1997)<br>(N=24) | .87 | NR | .84 (Total)<br>.91 (Obs.)<br>.66 (Comp) | NR | .62 (LOI-CV-20)<br>.34 (CDI)<br>.37 (RMACS) | NR |

Note. LOI-CV (44 and 20)=Leyton Obsessional Inventory- Child's Version (44 item and 20 item). NR=Not Reported; NA= Not Applicable; Y= Yes scores on LOI-CV, R=Resistance scores on LOI-CV, I=Interference scores on LOI-CV. OCR=Obsessive Compulsive Rating (Rapoport et al., 1980); CPRS-OC=Comprehensive Psychopathological Rating Scale-Obsessive Compulsive Subscale (Thoren et al., 1980); NIMH-OC= NIMH-Obsessive-Compulsive Scale (Rapoport et al, unpublished); NIMH-G=NIMH-Global scale (Murphy et al., 1982). CY-BOCS = Child's Yale-Brown Obsessive Compulsive Scale. Obs=Obsessions. Comp=Compulsions. CDI=Child Depression Inventory. RMACS= Revised Child Manifest Anxiety Scale.

### The Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS)

Like the LOI-CV, the CY-BOCS (Goodman, Rasmussen, Price, Riddle, & Rapoport, 1990) is a downward extension of the adult version, Y-BOCS. The CY-BOCS is a 19-item, clinician-administered scale, and is the most widely used outcome measure for childhood onset OCD. The first 10 items assess core OCD symptomatology (Items 1-5: obsessions; Items 6-10: compulsions), and comprise the total score. The remaining items asses associated features of OCD (5 items), global severity and improvement (2 items), reliability (1 item), and insight (1 item). All items are rated on a 5-point, Likert-type scale.

As is true for the LOI-CV (both 44-item and 20-item versions), there has been a paucity of research that has systematically investigated the technical properties of the CY-BOCS. This fact is surprising given its widespread use in drug outcome studies (Goodman & Price, 1992). A literature search was conducted through the Obsessive Compulsive Information center at the Dean Foundation located in Madison, Wisconsin by using the general descriptor, "Child Y-BOCS." This search revealed 24 articles. Only three of these 24 articles pertained to technical test information. Riddle et al. (1993) used 19 taped interviews and two blind raters to calculate inter-rater reliability, intraclass correlation coefficients, and a one-way ANOVA. Unfortunately, these statistics were not reported in this abstract, and it is unclear if a paper with these results was ever published. These authors indicated that validity measures would be computed in the future. Hanna (1995) reported validity measures of the CY-BOCS for a sample of 31 clinically referred children and adolescents with OCD. A Pearson product moment correlation of .93 was found for the CY-BOCS and the NIMH-G. Scahill and collegues (1997) investigated the reliability and validity of the CY-BOCS. Interrater reliability was established through 24 interviews which were videotaped and scored by four raters. Reportedly, the CY-BOCS was sensitive to change for 11 out of the 17 subjects who participated in a controlled drug study. The CY-BOCS detected change for 7 out of the 9 participants receiving the treatment and 4 out of the 8 in the placebo group. Given that the technical properties of the CY-BOCS are relatively untested, it is difficult to make a judgment about the reliability, validity, and utility of this instrument. (see Table 1 for technical information on CY-BOCS.)

Despite the scarcity of technical information, the advantage of the CY-BOCS is its ability to assess several dimensions of OCD symptoms such as severity, interference, distress, degree of control, frequency, insight, avoidance, and improvement. The greatest concern, however, is that the authors did not retain the symptom list from the adult version; therefore, a clinician might inadvertently miss a key OCD symptom. Stallings and March

(1995) conclude that there may be no clear advantage of the CY-BOCS over the Y-BOCS.

The LOI-CV and the CY-BOCS are probably the most common rating scales used for the assessment of OCD. These scales measure various characteristics of OCD and might be useful in certain situations. However, they share several limitations such as lack of adequate standardization as well as limited reliability and validity (specifically content validity and accuracy). Moreover, the limited and untested support for these devices attests to the importance of additional empirical work to extend the existing literature base and for alternative scale development.

## Direct Observation Assessment

Although there are no systematic direct observation measures currently used on a regular basis, this assessment method should be considered. Direct observation strategies afford the observer the framework necessary to record behavior as it naturally occurs. There are several different methods for obtaining direct observations (i.e., frequency, duration, event, latency, and narrative). The chosen method will depend on the purpose of assessment. For example, if an observer is interested in the number of times the target child raises her/his hand, an event recording will likely be used given the discrete nature of hand raising. Direct observations are also useful for the functional analysis of behavior in terms of antecedent, consequent, and sequential events.

Additionally, direct observations provide a more complete understanding of the context and content of symptoms (Francis & Gragg, 1996) This method is the most direct method of behavioral assessment (Cone, 1978) and has the distinct advantage of having ecological validity. Currently, there are no systematic direct observation measures for diagnosis, treatment planning, or treatment monitoring of OCD. Work in this area is noticeably absent given researchers' pleas for direct observational measures (Berg, 1989; Francis & Gragg, 1996; Henin & Kendall, 1997; Taylor, 1995).

## Reliability and Accuracy of Ratings and Direct Observation Measures

A considerable empirical body of literature exists concerning the reliability of self and other ratings. The literature on the reliability and accuracy of ratings and direct observation measures for children's social behavior is reported here in an effort to demonstrate how the accuracy-reliability paradigm has been investigated. No accuracy information has been used to validate OCD assessment measures, and reliability data on the two most frequently used OCD measures can be found in Table 1. Thus, the literature on children's social behavior can be offered to contextualize the proposed study.

Achenbach, McConaughy, and Howell (1987) authored a seminal review article regarding cross-informant agreement. Achenbach et al. used the meta-analysis technique to assess the overall level of consistency between the ratings of different informants. Their cross informant correlations were not designed to assess reliability because target behavior was not constant across all studies; instead, they attenuated several studies to determine the typical level of consistency between different informants. Their meta-analysis included 269 samples in 119 studies where the participants' age ranged from 1½ to 19 years old. The results of their meta-analysis revealed several noteworthy findings. First, the mean correlations for pairs of informants having similar relations to the participants (i.e., pairs of teachers, observers, parents, and mental health workers) ranged from .54 for pairs of mental health workers to .64 for pairs of teachers. The mean correlation for observer-observer was .57. Second, for pairs of informants with different relationships to the participants (i.e., parent-teacher; parent-mental health worker; parent-observer; teacher/mental health worker; and teacher/observer), the teacher-observer pair demonstrated the highest mean correlation, .42. Finally, all mean correlations between the different types of informants were statistically significant lending support to the idea that there is a certain degree of consistency among raters. It can be concluded that there is higher consistency between informants in similar roles than informants in different roles and different situations. Due to the situational specificity of behaviors, correlating scores of multiple raters will appear to reduce the interrater reliability. This finding, however, should not preclude the use of multiple informants since multiple informants across multiple situations often are needed to gain a more complete topographical understanding of a behavior.

A critical study that examined the accuracy and reliability among observations and ratings was conducted by Robertson (1993). This investigator used a videotape of a Head Start preschooler as an incontrovertible index. The child's teacher viewed the videotape and concluded that the videotape was representative of the child's behavior. Two experienced observers also came to 100% agreement in regard to the target child's behavior on the videotape, and thus, were considered to render an "expert" criterion to which other assessment results were compared. Hence, a standard of accuracy was established. Robertson also developed an event recording observation system to be used as a companion to the Social Skills Rating System - Teacher form (SSRS-T). The expert observation was used as the criterion against which the observations of four, trained graduate students were compared. The design of her study was referred to as the video-accuracy paradigm (Robertson, 1993). More precisely, this design can be referred to as an accuracy-reliability paradigm.

Robertson found that there was a moderate correlation between teacher and observers' ratings (M= 58) Observers achieved a high interrater agreement among themselves (M= 66), and the average agreement between expert and observers, which essentially assessed accuracy, was .89. Low to moderate correlations (mean r = .38), however, were obtained between observers' observations and the teacher's rating. Comparing observers' ratings yielded a mean correlation of .60 The correlations reported in the Robertson study were much higher than those reported in the Achenbach et al. (1987) meta-analysis. Robertson speculated that this discrepancy could have been due to the criterion mastery training approach used in her investigation. In essence, Robertson's study suggests that training individuals as observers may account for more accurate and reliable ratings.

Racine (1994) sought to replicate Robertson's (1993) study by incorporating additional design features to test the accuracy and reliability of direct observations and ratings of social behavior. Unlike Robertson's study, this investigation incorporated the following: (a) experts' ratings in addition to experts' observations; (b) observers only (not actual teachers); (c) an increased sample size of observers (from N = 4 to N = 10); and (d) two videotaped case vignettes of two different preschool age children. Racine found moderate to high correlations between observers' SSRS ratings, which were comparable to Robertson's results. Additionally, moderately high percent agreements and Pearson correlations between observers' direct observations were found on both videotapes.

Upon the suggestion of Racine (1994), Seymour and Kratochwill (1998) extended the accuracy-reliability paradigm to scale development. Seymour developed the Selective Mutism Analysis Scale (SMAS) and the Selective Mutism Observation Protocol (SMOP) by using expert consensus as the incontrovertible index. Graduate students (*N*=8) completed the SMAS and the SMOP while watching the scripted, analog videotape. Unlike the Robertson (1993) and Racine (1994) studies, experts and graduate students were allowed to stop and rewind the videotapes as many times as they wanted. In terms of accuracy, Seymour and Kratochwill documented moderate correlations between graduate students' performance on the SMAS and the SMOP (*M* *K* = .67and .64, respectively). With respect to reliability, graduate students demonstrated moderate correlations between pairs of observers on the SMOP (*M* *r* = .54) and pairs of raters on the SMAS (*M* *r* = .56).

## Relationship Between Ratings and Direct Observations

Recently, several investigators interested in the development of behavior rating scales have reported data concerning the relationship between ratings of children's behavior and direct observations of behavior. None of these studies, however, has assessed this relationship using similar informants except those done by Robertson (1993) and Racine (1994). Three studies investigating cross-informant agreement between direct observations and ratings are reviewed here.

Reed and Edelbrock (1983) conducted a study that sought to validate an observation system. *The Child Behavior Checklist-Direct Observation Form* (CBCL-DOF. Achenbach & Edelbrock, 1983), which was developed to specifically complement the *Child Behavior Checklist* (CBCL. Achenbach & Edelbrock. 1983). The CBCL-DOF contains 96 problem behavior items as well as a measure of on-task behavior. For the problem behavior measure, the observer watches the target child for 10 minutes and then completes the CBCL-DOF by assessing each item on a 0-1-2-3 response scale. Even though the CBCL-DOF is called an observation system, the CBCL-DOF could more appropriately be termed a checklist rather than an event recording procedure. The on-task score is calculated somewhat differently from the problem behavior scale. After one minute of observation, the observer takes a five second interval to assess if the target child is on task. If the child is on task then she/he receives a score of 1. This continues up to 10 minutes so a score between 0-10 is derived.

In this study of boys, ages 6- to 11- years-old (*N*=25), referred for emotional/behavior problems, direct observations were gathered by two graduate students. The boys, were observed six times across a two-week period. The teacher also completed the CBCL for each child. Reed and Edelbrock found that direct observations of problem behavior correlated positively with teacher ratings of problem behavior ($r$ = .44 and .37, respectively); negatively with school performance ($r$ = -.60 and -.49) and negatively for adaptive functioning ($r$ = -.69 and -.61). In addition, inter-observer agreements were calculated for total behavior problem and on-task scores, and these were: .91 and .83, respectively. Inter-observer reliabilities for each of the six 10 minute sessions were also computed. The inter-observer reliability for the behavior problem score was an average of .85 (Range = .81 to .92) and a .71 (Range = .56 to .87) for the on-task score.

Elliott et al. (1988) compared teacher and observers' ratings on an experimental version of the SSRS-T to classroom observations. The observational coding system was based on five classroom behaviors: (a) working productively, (b) desirable peer interaction, (c) undesirable peer interaction, (d) passive appropriate, and (e) passive inappropriate. These researchers discovered different patterns of correlation for observer and teacher ratings on the SSRS-T and observed classroom behaviors; for example, observer ratings on Academic Performance were not strongly correlated with work engagement ($r$ = .12). Interestingly, teacher ratings on the Academic Performance subscale were correlated negatively with the percentage of intervals recorded in actual academic work ($r$ = -.42). The authors assert that the different patterns of correlations could be due to the low frequency of some behavior categories. These different patterns of correlations between observations and ratings also appear to support the idea that ratings of behaviors may be based more upon isolated incidents of student behavior rather than on averages of student behavior over time (Elliott et al., 1988). For the most part, teachers' and observers' ratings on the SSRS-T were moderately correlated with certain categories of the directly observed classroom behaviors. Teachers' ratings on the Social Initiation and Cooperation factors were correlated with observed desirable social interaction (r = .45 and .55, respectively). In contrast to these findings, observers' SSRS-T ratings on the Social Initiation and Cooperation factors were not correlated significantly with observed desirable social interaction (r = .16 and .17, respectively).

Merrell (1993) found similar results to that of Elliott et al. (1988) in regards to teacher ratings and observer observations. Merrell examined the development, standardization, and validation of the School Social Behavior Scales (SSBS). One of the criterion-related validity studies that Merrell conducted was correlating the SSBS with the CBCL-DOF. The correlations between the teacher ratings on the SSBS and the observations by two trained graduate students on the overall CBCL-DOF were weak to moderate (r = -.06 to -.39) on the Social Competence scale and weak (r = .01 to .27) on the Problem Behavior scale. Low to moderate correlations were discovered for on-task ratings, which ranged from .26 to .52 on the Social Competence scale and -.22 to -.37 for the Problem Behavior scale. Merrell concluded that it was not surprising that these relationships were weak to moderate, at best, considering that the two instruments were not developed to be companions and the CBCL-DOF is really more of a checklist than a direct observation measure. Therefore, the two assessment techniques are likely valid for different purposes, and measure different constructs. Because of the low agreement between observations and ratings, Merrell stressed the need to include both observations and ratings in behavioral assessments of children.

Robertson (1993) found a moderate correlation of .60 between observers' SSRS rating and their own direct observation (N = 4). Similarly, Racine (1994) found a moderate correlation of .59 between observers' SSRS rating and their own direct observation (N = 10) on the same videotape that Robertson used in her study. Seymour and Kratochwill (1998) also found moderate correlations between an observer's SMOP and SMAS for graduate students (M r = .58) and for experts (M r = .62)These correlations are somewhat higher than those reported by Merrell (1993), Elliott et al. (1988), and Reed and Edelbrock (1983). Racine (1994) concluded that weaker correlations between observations and ratings have been found when rating and observation instruments have been developed independently of one another, and direct observations have been conducted in naturalistic settings. Those instruments that have been developed together and have used videotaped behavior, however, have resulted in higher correlations.

Again, this literature serves to contextualize the accuracy-reliability paradigm and illustrates how it can be extended into the domain of childhood disorders, such as OCD. Although only one investigator at this point in time has used this framework for scale development, the accuracy-reliability paradigm for scale development shows promise for the development of future OCD assessment measures.

## Summary and Predictions

Given that current OCD assessment measures have not documented appropriate validity indices for behavioral measures, the utility of these measures is questionable. Therefore, there was a need to develop new instruments, such as a rating scale and direct observation measure in accordance with Foster and Cone's (1995) concept of representational validity (i.e., content validity and accuracy) for behavioral measures. Hence, the purpose of this study was to develop a rating scale and a direct observation measure to monitor OCD treatment (psychosocial and/or psychopharmacological). The accuracy-reliability framework was used to guide development and validation of the measures. Given the existing research base, the following predictions were offered and tested:

*Prediction 1.* The OCD field experts' observations and ratings would be in high agreement ($r$ = .70 to .90) with the experts' observations and ratings from UW-Madison (i.e., the performance standard or incontrovertible index).

*Prediction 2.* Trained graduate student raters would have moderately high inter-rater reliability ($r$ = .50 to .80) on their ratings.

*Prediction 3.* Trained graduate student observers would have high inter-observer agreement ($r$ = .60 to .90) when comparing their direct observations.

*Prediction 4.* Trained graduate students' observations and ratings would be highly accurate when compared to the performance standard's observations and ratings.

*Prediction 5.* An observer's observation and his/her own ratings would correlate moderately ($r$ = .30 to .60).

*Prediction 6.* The new OCD rating scale and direct observation system would be sensitive to behavior change as reflected on videotape.

# Chapter 3

## Method

### Participants

The participants in this study were: (1) school-age actors/actresses performing the roles of high school students (N=9); (2) a school-age actress portraying an adolescent with OCD (N= 1); (3) an adult actor portraying the teacher (N=1). (4) trained graduate student observers (N=6) from the University of Wisconsin-Madison's Department of Educational Psychology; and (5) experts (N = 9) in the field of OCD. All participants were recruited by direct verbal contact and/or by mail. Participation was voluntary. Consent letters informing the participants of the purpose and nature of the study were provided and signed by each participant (See Appendix A). Parental consent was obtained for minors.

### Materials

Two 25-minute videotapes, an OCD rating scale, and an OCD direct observation system were the materials used in this investigation.

New instruments. The most widely used outcome measures of childhood OCD, the LOI-CV and the CY-BOCS, were used as templates along with clinical expertise to develop an individually tailored OCD Rating Scale (OCD-RS) and a companion direct observation measure, OCD Direct

Observation System (OCD-DOS), that can by used in the monitoring of any OCD intervention.

Given that this study used scripted, videotaped classroom scenarios, this researcher had the ability to choose and operationally define those OCD symptoms most likely to be seen in adolescents. Criteria for selection of target behaviors was determined by those OCD symptoms most commonly found in children and adolescents as reported in the literature (Rappoport, 1992). In addition to literature review, practitioner (i.e., psychiatrists, psychologists, and social workers) perspectives also were solicited when choosing four target behaviors for the direct observation measure and rating scale. The four target behaviors chosen were reassurance seeking, writing rituals, repeating, and, symmetry. See Appendix B for opperational definitions of target behaviors.

The OCD-RS was designed to measure retrospectively how often a student exhibits specific, observable OCD behaviors in the classroom, home, and/or community. The OCD-RS also assesses how these specific behaviors interfere with a student's success in a particular environment. The interference dimension was created because it is congruent with the DSM-IV criteria for OCD and it gives the evaluator an idea as to which behavior needs the most intervention. The OCD-RS can be analyzed by item (target behavior) or across the entire scale. For purposes of this study, four target behaviors were chosen. These target behaviors were rated on a frequency dimension ( 0=Never, 1=Sometimes, and 2=Very Often) and an interference dimension (0=Never, 1=Sometimes, and 2=Most of the Time).

The OCD-DOS was designed as a structured direct behavior observation system for observing the compulsive behaviors of children and adolescents who have OCD. This observation system only focuses on the overt, compulsive behaviors commonly associated with OCD. The OCD-DOS uses an event recording procedure to measure the frequency of target behaviors. The OCD-DOS requires the observer to record the target behavior with a tally mark for each occurrence of the behavior during the observation period. A copy of the OCD-RS and OCD-DOS may be found in Appendix B along with the operational definitions of the four target behaviors used in this study.

Videotape(s). Two 40-minute, 10th grade language arts class videotapes were created for the purpose of obtaining a naturalistic sample of an adolescent's observable OCD behavior in a school classroom. The videotapes were strategically designed to capture the observable OCD behaviors along a continuum (i.e., Language Arts Class #1 and Language Arts Class #2). Hence, reassurance seeking, writing, and repeating occurred with more frequency in videotape #1 than in videotape #2, and there were more instances of symmetry in videotape #2 than in videotape #1.

Adolescent actors were trained prior to the actual taping of the classroom scenarios. These training sessions were held on two separate occasions for two hours each. The target adolescent who portrayed the four symptoms of OCD went through a more intense training that included one-on-one training sessions (N=4) with this researcher and extensive reading in the area of OCD. The taping of these two classroom situations occurred on a single day at the University of Wisconsin-School of Nursing Audio-Visual center. The videotape was then viewed by two experts in the field of OCD and determined to be naturalistic. These two videotapes were used as the incontrovertible indices in this study.

Procedure

Incontrovertible Index. Two experienced observers and OCD experts reviewed the two videotapes several times and came to 100% agreement on the target adolescent's behavior as demonstrated on each videotape. This expert review was facilitated by the trainer's general idea about when the target behavior occurred due to the scripted nature of the videotape. These OCD experts consensus recording on the OCD-DOS became the performance standard, or incontrovertible index. These experts then completed the OCD-RS and came to consensus on their ratings; hence, a performance standard for the OCD-RS was created. The performance standard on the OCD-DOS and OCD-RS was used as the accuracy criterion against which the graduate students' and OCD experts' observations and ratings were compared.

Observer training. The graduate student observers were trained as a group for 90 minutes on each of two occasions. An observation manual was created to facilitate mastery and understanding of OCD and the OCD-DOS. Careful discussion of the target behaviors took place in an effort to clarify the behavioral definitions. Included in the manual was a behavior codes quiz in which the observers had to attain 100% mastery before going on to the next phase of training. After this criterion was met, the OCD-DOS was introduced. Next, the observers were trained using videotapes of different children and adolescents exhibiting OCD symptoms. The observers were trained by the researcher until they each reached a mastery criterion of 85% agreement.

Rater training. Like the OCD-DOS, a manual was used in an effort to explain and facilitate use of the OCD-RS. Actual rater training consisted of familiarizing each rater with the OCD-RS and discussing different strategies that can be used for the rating process. For example, some raters watch for the whole time and then rate, while others retain a certain score mentally and adjust the score depending on what is observed. Raters practiced rating after they viewed the videotapes of different children and adolescents exhibiting OCD symptoms. A complete copy of the manual used in training can be found in Appendix C.

## Data Collection

Data collection began a week after formal training took place. The six graduate students were randomly assigned to one of two groups. Group 1 (n=3) saw videotape #1 followed by videotape #2. These three observers did the following: (a) watched videotape #1 for 25-minutes while recording on the OCD-DOS; and then (b) watched videotape #1 again and then rated the four behaviors using the OCD-RS. This same sequence was repeated using videotape #2.

The sequence for Group 2 (n=3) was reversed to negate any order effect that might occur due to taped content. These three observers did the following: (a) watched videotape #2 for 25-minutes and then rated the behavior using the OCD-RS, and (b) watched videotape #2 again, this time recording the behaviors on the OCD-DOS. This same sequence was then continued with videotape #1. An interval of one week separated each task for both Groups 1 and 2. The entire data collection process utilized with graduate students is summarized in Table 2.

A total of 12 OCD experts were chosen and contacted for their extensive writing and research in the area of OCD or by their clinical work with children and adolescents who have OCD. These names and addresses were obtained through the Obsessive-Compulsive Information Center in

Table 2

Data Collection Procedure Used with the Graduate Students After Criterion-Mastery Training

| | Videotape #1 (Language Arts Class #1) | | Videotape #2 Language Arts Class # 2 | |
|---|---|---|---|---|
| 1 | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating |
| 2 | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating |
| 3 | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating |
| PS | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating |

| | Videotape #2 (Language Arts Class #2) | | Videotape #1 (Language Arts Class #1) | |
|---|---|---|---|---|
| 4 | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS |
| 5 | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS |
| 6 | OCD-RS Rating | Observe Using OCD-DOS | OCD-RS Rating | Observe Using OCD-DOS |

Note. A week elapsed between each task. Group 1 = Observers 1-3. Group 2= Observers 4-6. PS = Performance standard.

Madison, Wisconsin. These 12 OCD field experts were sent a manual with explicit instructions on how to use the observation system and rating scale as well as a videotape with both classroom scenarios. They were asked to read the manual, and complete the instruments as they were instructed, and in turn, they would be paid a stipend of $500. They were to asked to complete the process just as Group 1 of the graduate student observers. That is, they would watch the first videotape while completing the OCD-DOS, and then complete the OCD-RS. They were asked to do the same for videotape #2. In addition, they were instructed not to view the videotape multiple times or stop and rewind. A week did not have to separate each task.

### Research Design and Data Analysis

The major dependent variables in this study were: (a) the frequency summaries for each target behavior on the OCD-DOS, and (b) the OCD-RS ratings produced by the six graduate student observers/raters and the nine OCD field experts. This study utilized a descriptive research design and analysis to examine accuracy and reliability of the OCD-RS and OCD-DOS.

Prediction 1. The OCD field experts' observations and ratings would be in high agreement ($r$ = .70 - .90) with the experts' observations and ratings from UW-Madison(i.e., the performance standard). This prediction was tested by computing the Root Mean Square for each target behavior.

*Prediction 2.* Trained graduate student raters would have moderately high inter-rater reliability ($r$ = .50 to .80) on their ratings of the OCD-RS. This prediction was tested by calculating percent agreement.

*Prediction 3.* Trained graduate student observers would have high inter-observer agreement ($r$ = .60 to .90) when comparing their direct observations using the OCD-DOS. This prediction was tested by using Root Mean Square.

*Prediction 4.* Trained graduate student observers' observations and ratings would be highly accurate when compared to the performance standard's observations and ratings. This prediction was tested by computing the Root Mean Square, percent agreement, and Pearson correlations.

*Prediction 5.* An observer's observation and his/her own ratings would correlate moderately ($r$ = .30 to .60). This prediction was tested by computing Pearson correlations between observers' observations and ratings.

*Prediction 6.* The OCD-RS and the OCD-DOS would be sensitive to behavior change as reflected on videotape. This prediction was tested by calculating the effect size between videotape #1 and videotape #2 for the four target behaviors.

# Chapter 4

## Results

Direct observations and ratings were gathered from six trained graduate students and nine OCD field experts who viewed the two 25-minute analog classroom situation videotapes (videotape #1 and #2). The design of this study was descriptive in nature. Descriptive statistics, percent agreement, Pearson correlations, and effect size measures were the main data analytic techniques used to test the six predictions.

When examining the rating scale agreement, exact percent agreement was used versus occurrence/non-occurrence of behavior percent agreement. Occurrence/non-occurrence agreement is an upper limit percent agreement because raters are said to agree if they say the behavior regardless of how many times they saw the behavior (i.e., if rater 1 rated Reassurance Seeking as a 2 (very often) and rater 2 rated Reassurance Seeking as a 1 (sometimes), they are said to agree; even though, there is a discrepancy in the amount of times they saw this behavior.) Since this is a liberal estimate of agreement, exact percent agreement was chosen because it provides a lower limit in that it is computed by looking at the exact frequency agreement (i.e., both raters rated the behavior as a '2,' '1,' or '0').

Exact percent agreement for direct observations was not calculated for pairs of observers across the total or by item due the continuous nature of the direct observations. In general, all graduate student observers and OCD experts agreed on whether they saw the behavior (occurrence/non-occurrence percent agreement). Across graduate student raters, only one graduate student rated as not observing the behavior when it actually occurred (videotape #1; SY). Across graduate observations, the same graduate student did not observe an occurrence of a target behavior when there actually was an occurrence (videotape #2, W).

All OCD experts agreed that they saw the four target behaviors on videotape #1 using the OCD-RS and the OCD-DOS. On videotape #2, however, four OCD experts (E2, E6, E7, & E9) indicated on the OCD-DOS that they saw repeating rituals when these behaviors actually did not occur. Four experts (E2, E5, E7, & E9) rated Repeating as "Sometimes" when it did not happen at all. With the exception of these four OCD experts on videotape #2, there were very few times that all the participants did not agree on whether a behavior occurred. Given this finding, exact percent agreement is used for the remainder of the results to evaluate reliability and accuracy.

Pearson correlations are reported for some predictions; however they should be interpreted with caution given the small number of items on the OCD-DOS and OCD-RS as well as the lack of repeated measures. The magnitude of the correlations was of greatest importance to this researcher; however, statistical tests of significance were reported for each correlation.

Given the nature of the data, most traditional psychometric statistics do not apply. Thus, most of the statistics calculated in this study were descriptive to assess the accuracy of the observations and ratings. It is this specificity that lends meaning to the results. Results regarding each prediction are reported separately. The raw data, upon which these results are based, are provided in Appendix C. Figures 2 and 3 provide a visual reference for the OCD experts' amd graduate students' mean ratings and observations on each target behavior as compared to the performance standard. Figures 4 through 7 illustrate individual OCD expert's and graduate student's ratings and observations as compared to the performance standard.

Prediction 1: Partially Supported

It was predicted that the OCD field experts' observations and ratings would be in high agreement with the experts' observations and ratings from UW-Madison. Evidence was found to support this prediction by computing Root Mean Square (RMS). RMS was performed because it demonstrates how far away the collective observations were from the performance standard. In other words, this statistic illustrates the variance around the UW-Madison performance standard (PS). The farther away from 0 (no

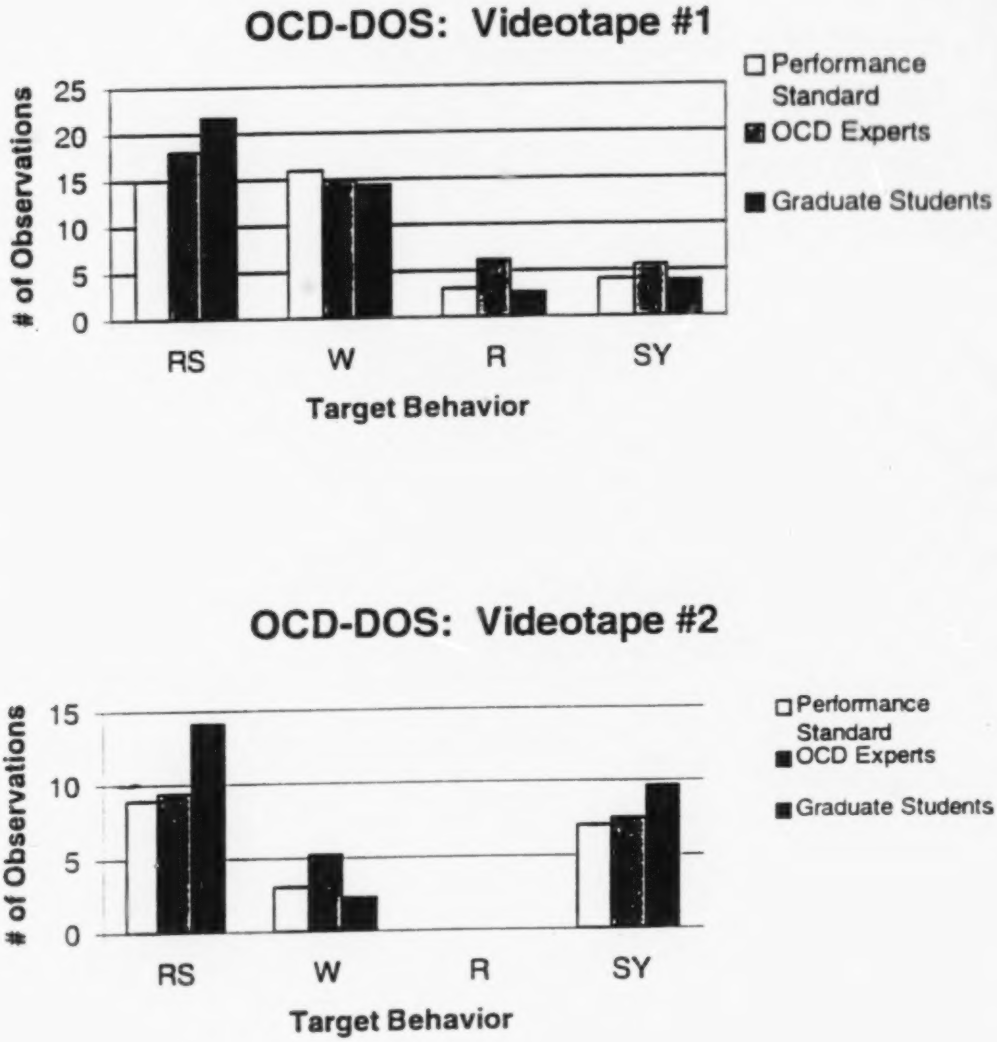


Figure 2. OCD experts' and graduate students' mean OCD-DOS observations compared to the performance standard's OCD-DOS observations for videotapes #1 and #2.

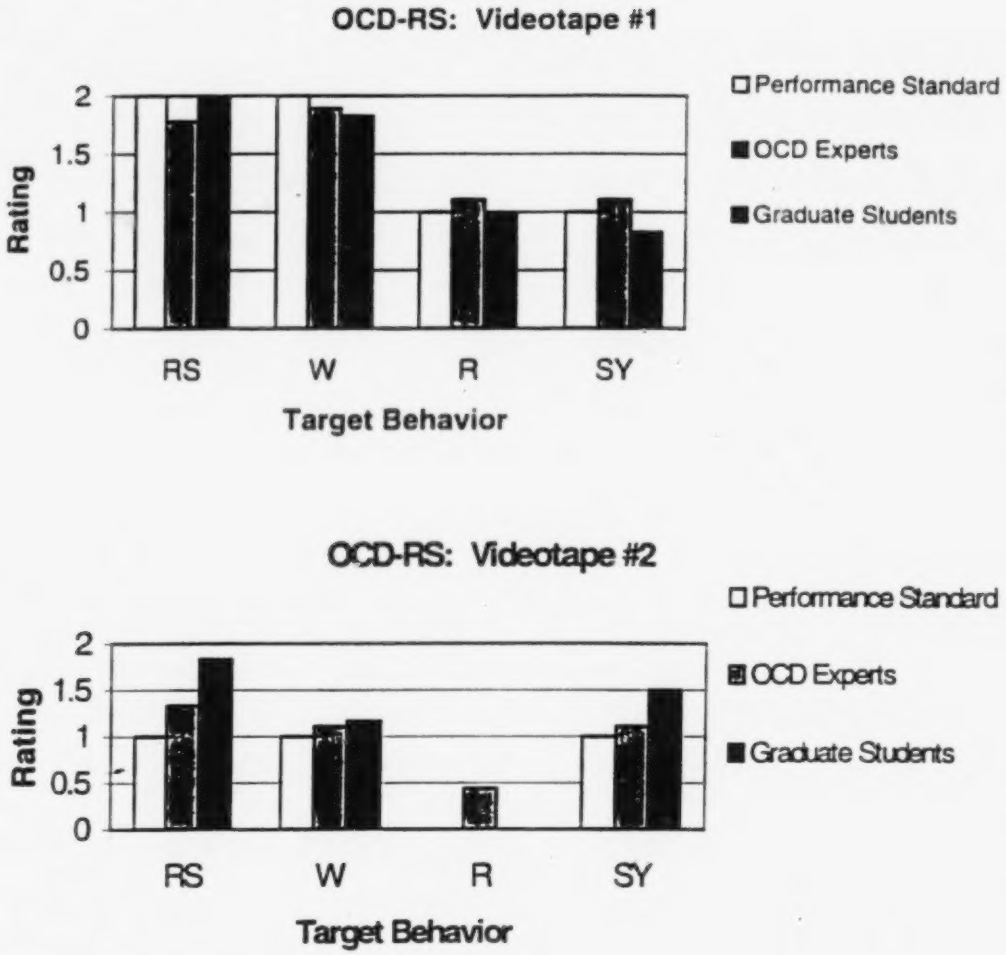


Figure 3. OCD experts' and graduate students' mean OCD-RS ratings compared to the performance standard's ratings for videotapes #1 and #2.

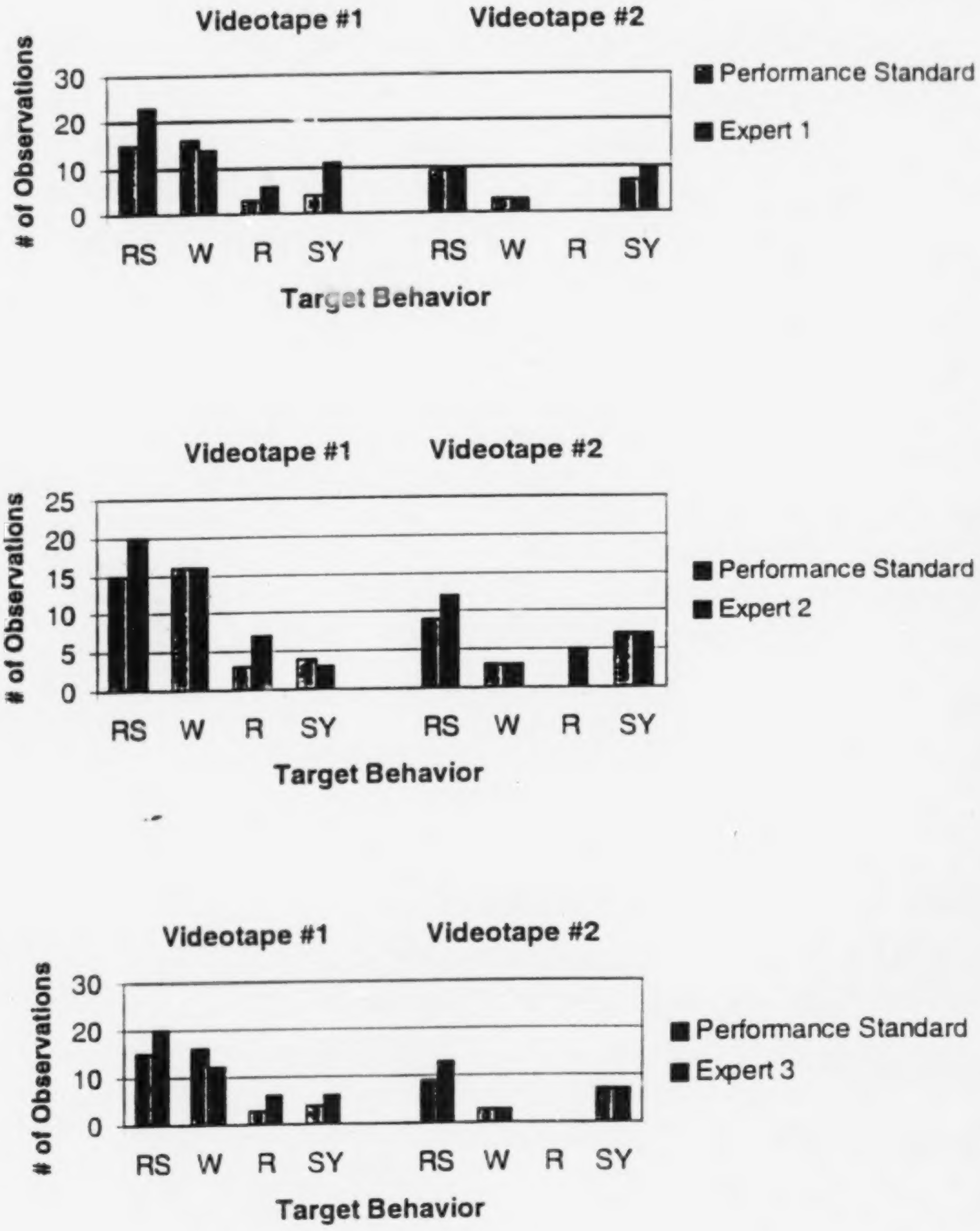


Figure 4. Individual OCD expert's observations on the OCD-DOS compared to the performance standard for videotape #1 and #2.

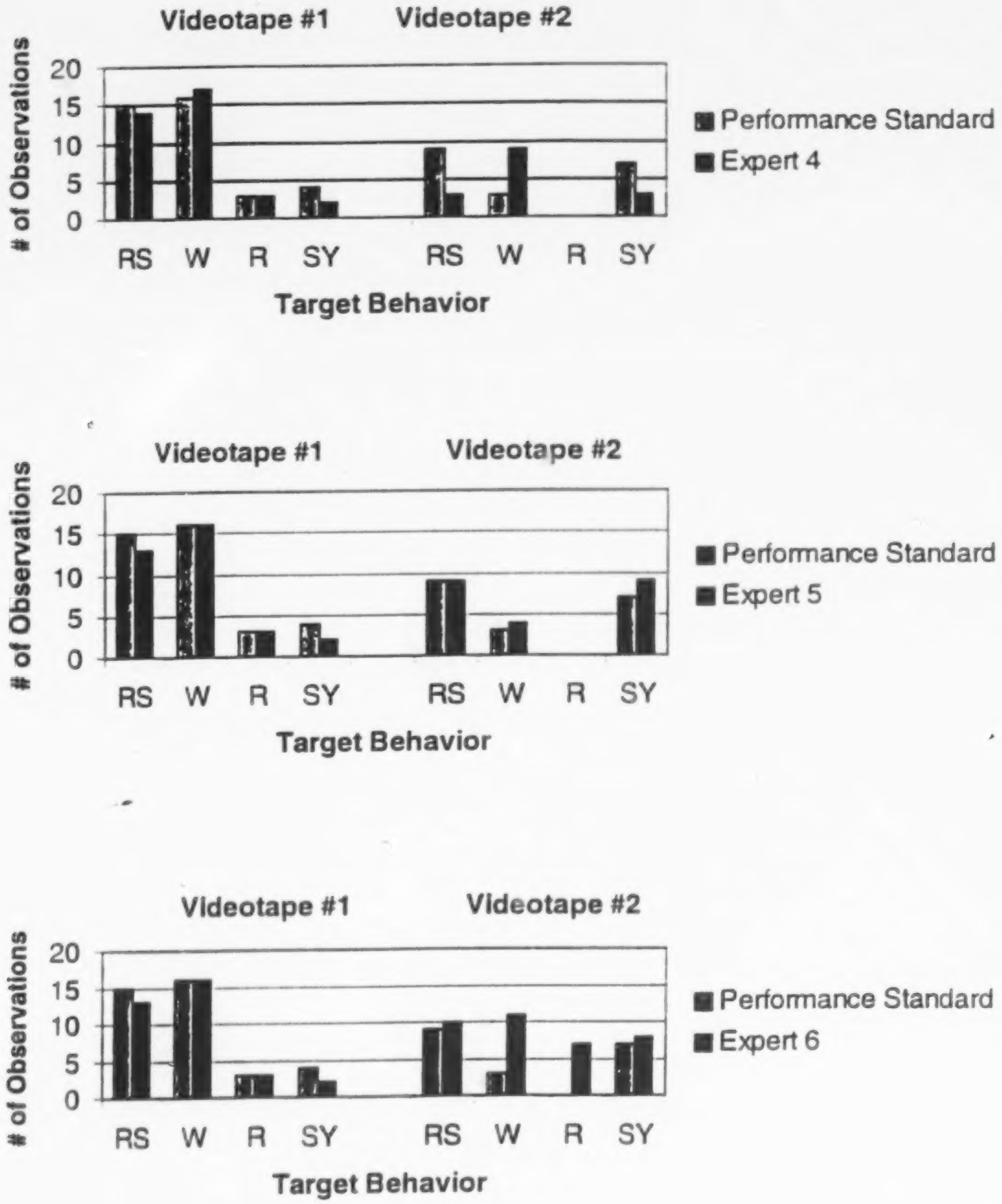
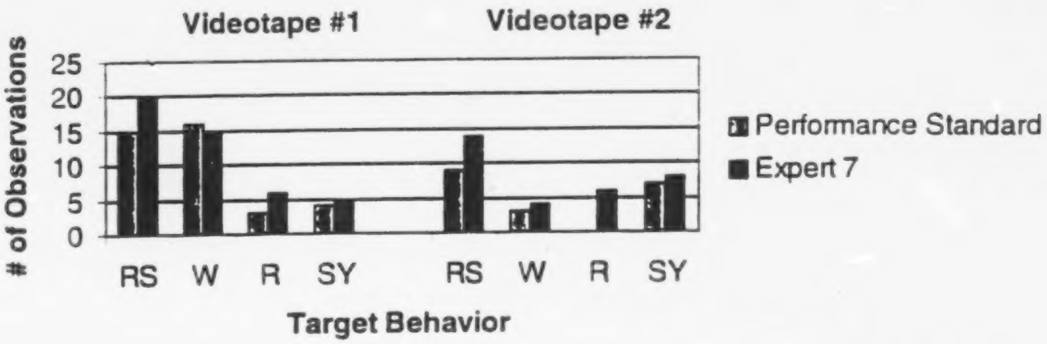
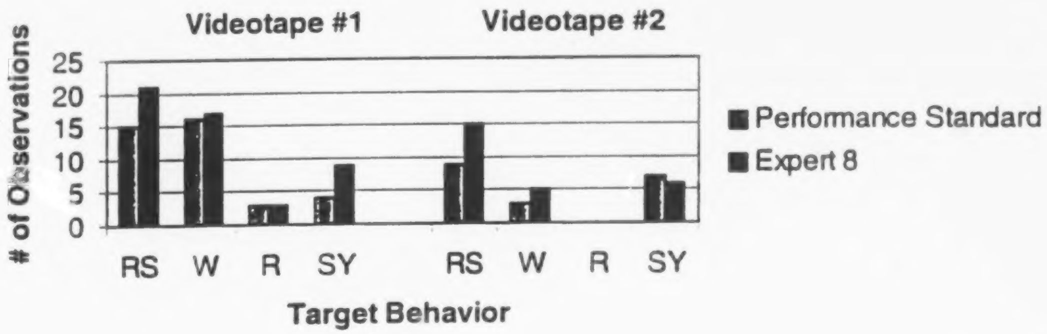


Figure 4 (Cont). Individual OCD expert's observations on the OCD-DOS compared to the performance standard for videotape #1 and #2.





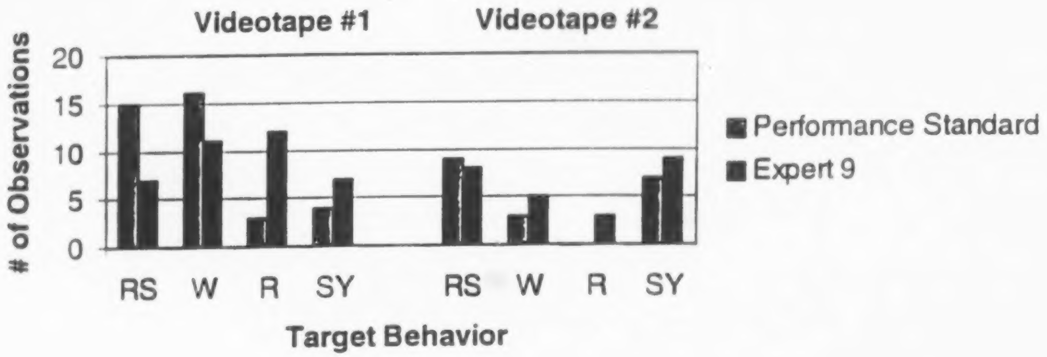


Figure 4 (Cont). Individual OCD expert's observations on the OCD-DOS compared to the performance standard for videotape #1 and #2.

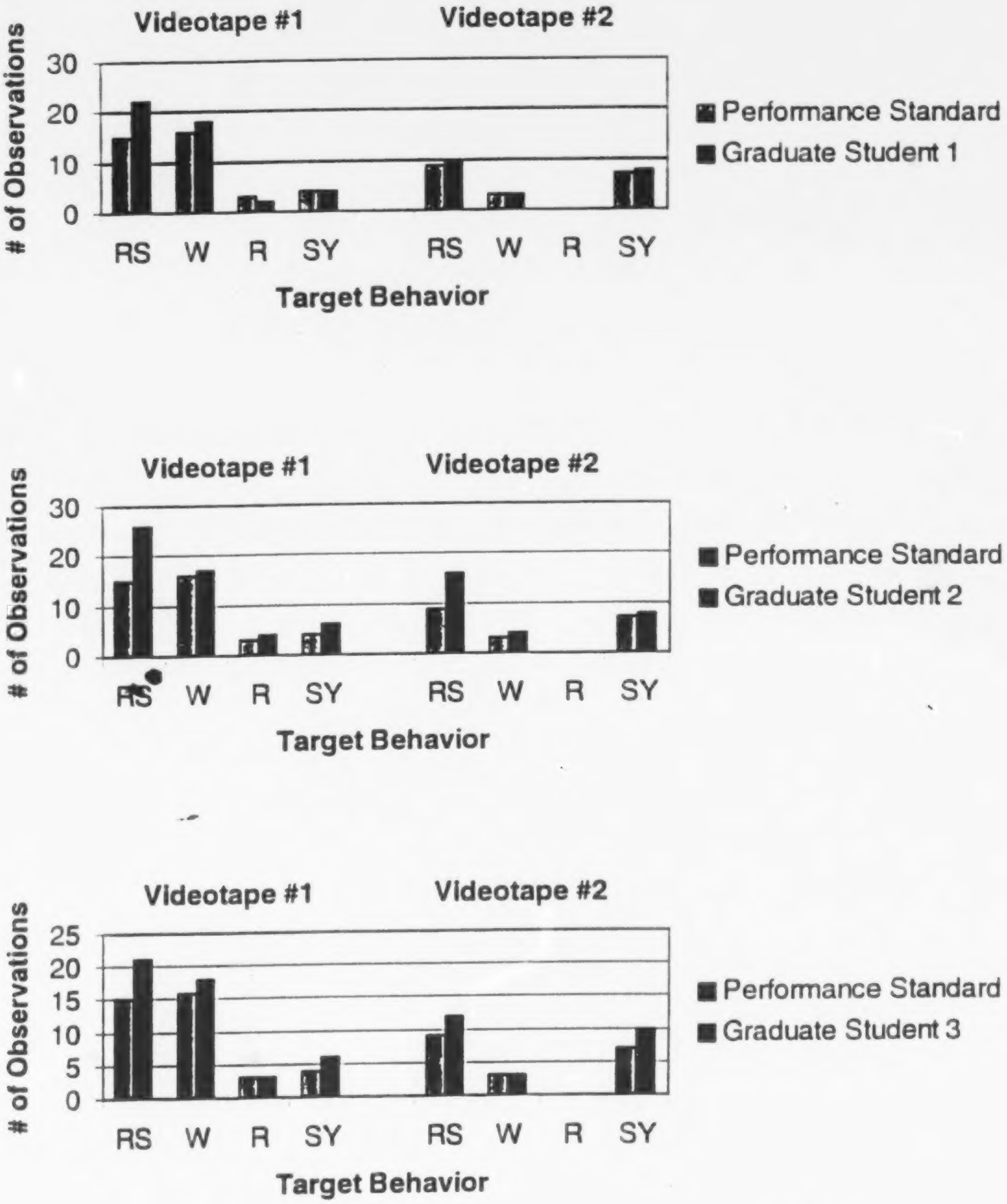


Figure 5. Individual graduate student's observations on the OCD-DOS compared to the performance standard for videotapes #1 and #2.

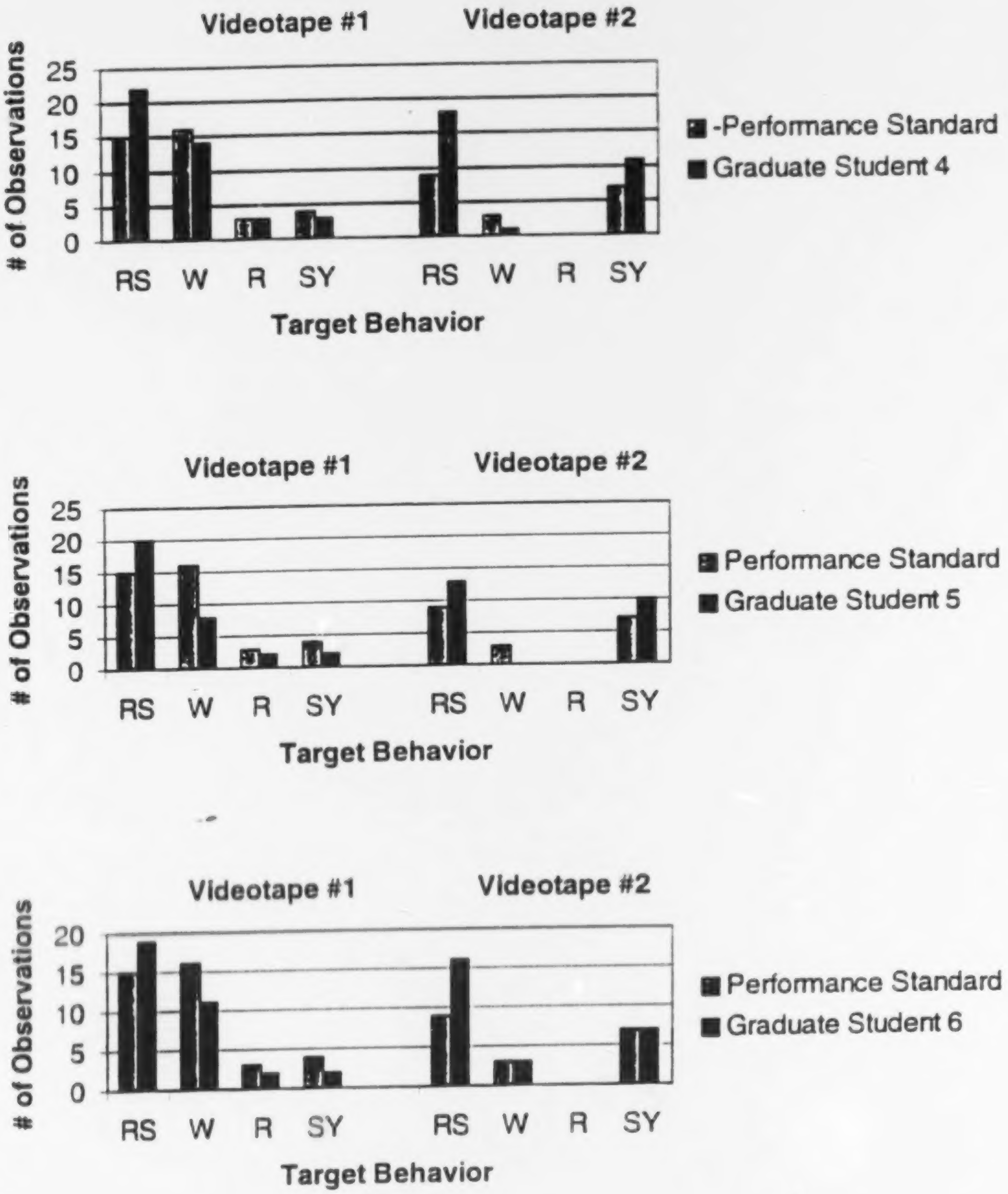


Figure 5 (Cont.). Individual graduate student's observations on the OCD-DOS compared to the performance standard for videotapes #1 and #2.

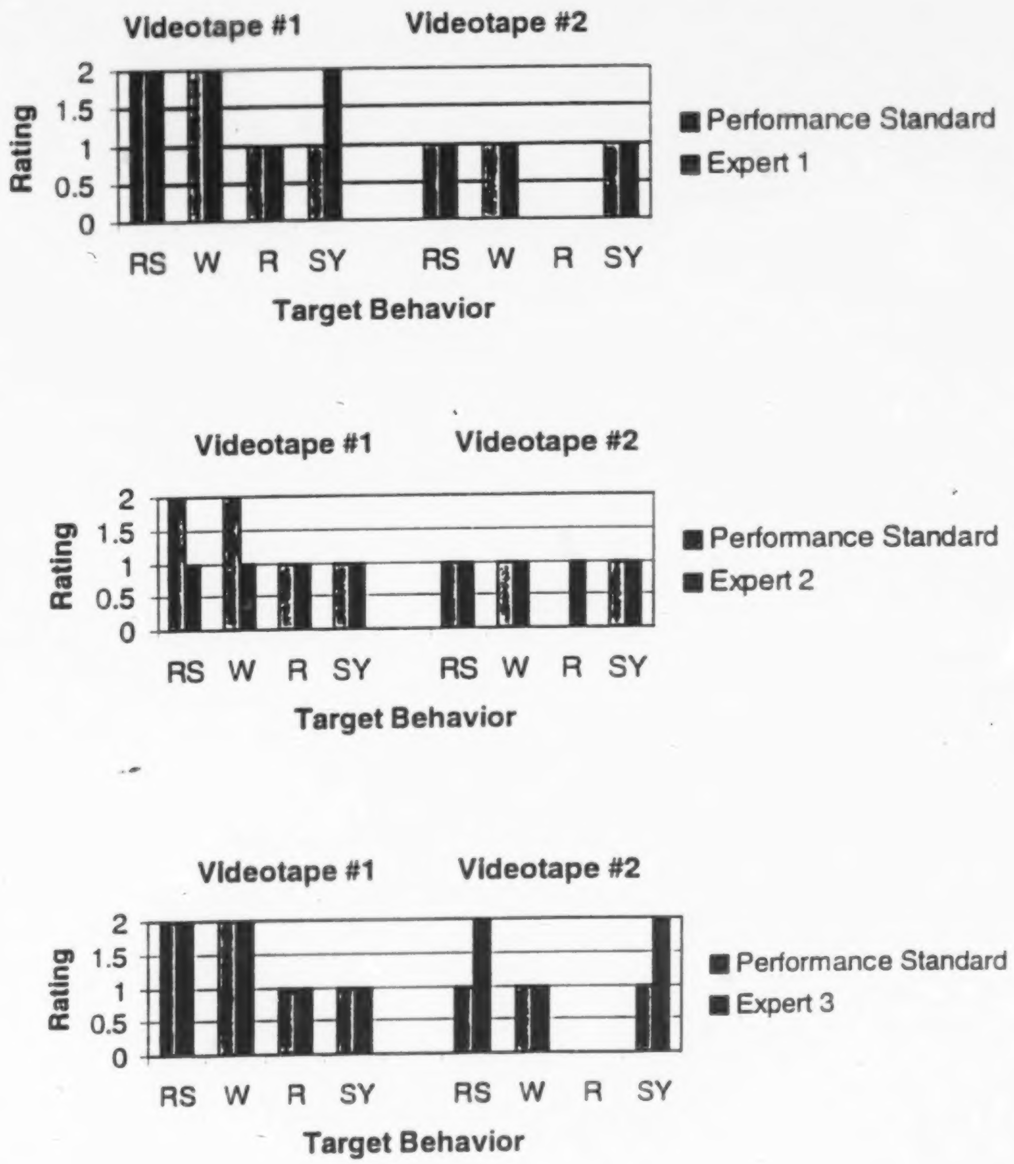


Figure 6. Individual OCD expert's OCD-RS ratings compared to the performance standard for videotapes #1 and #2.

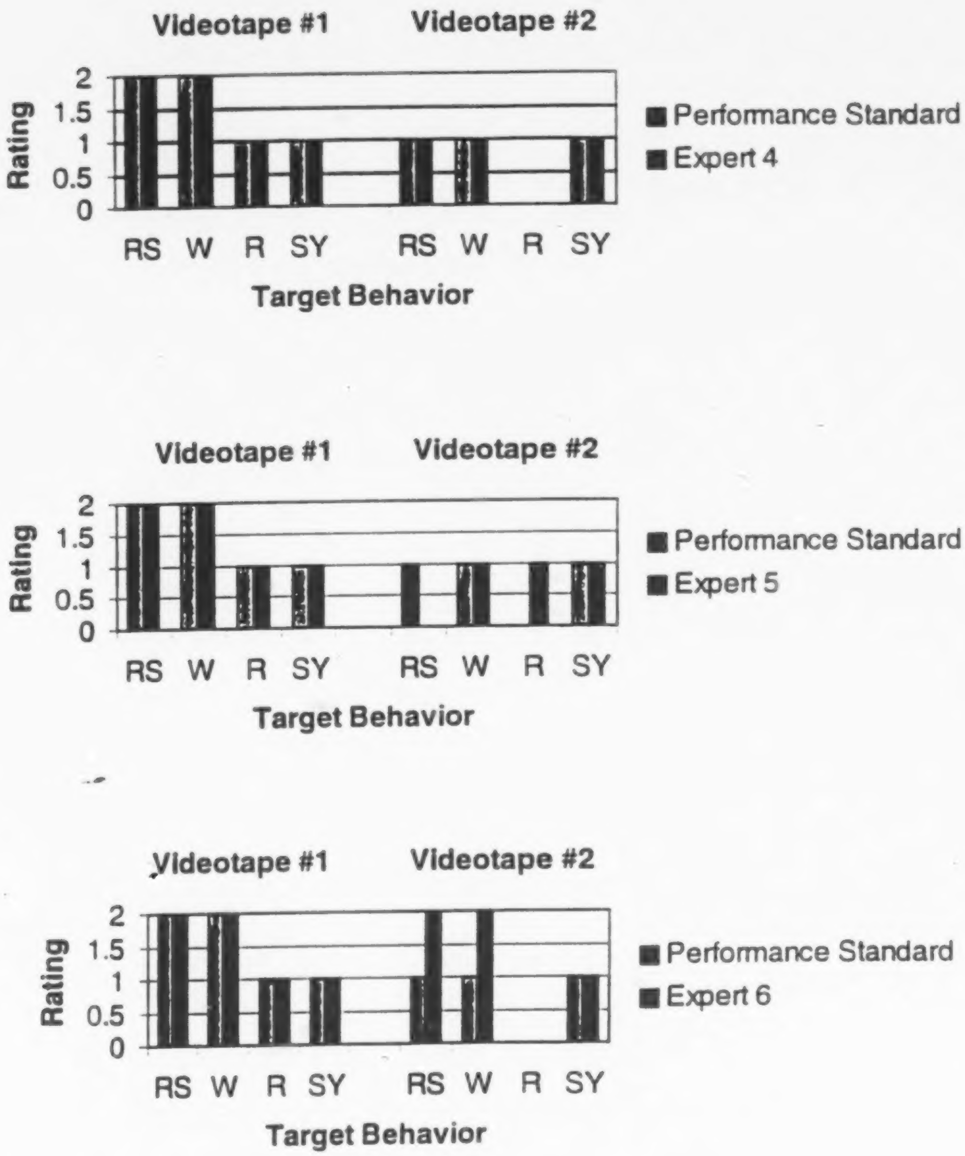


Figure 6 (Cont.). Individual OCD expert's OCD-RS ratings compared to the performance standard for videotapes #1 and #2.

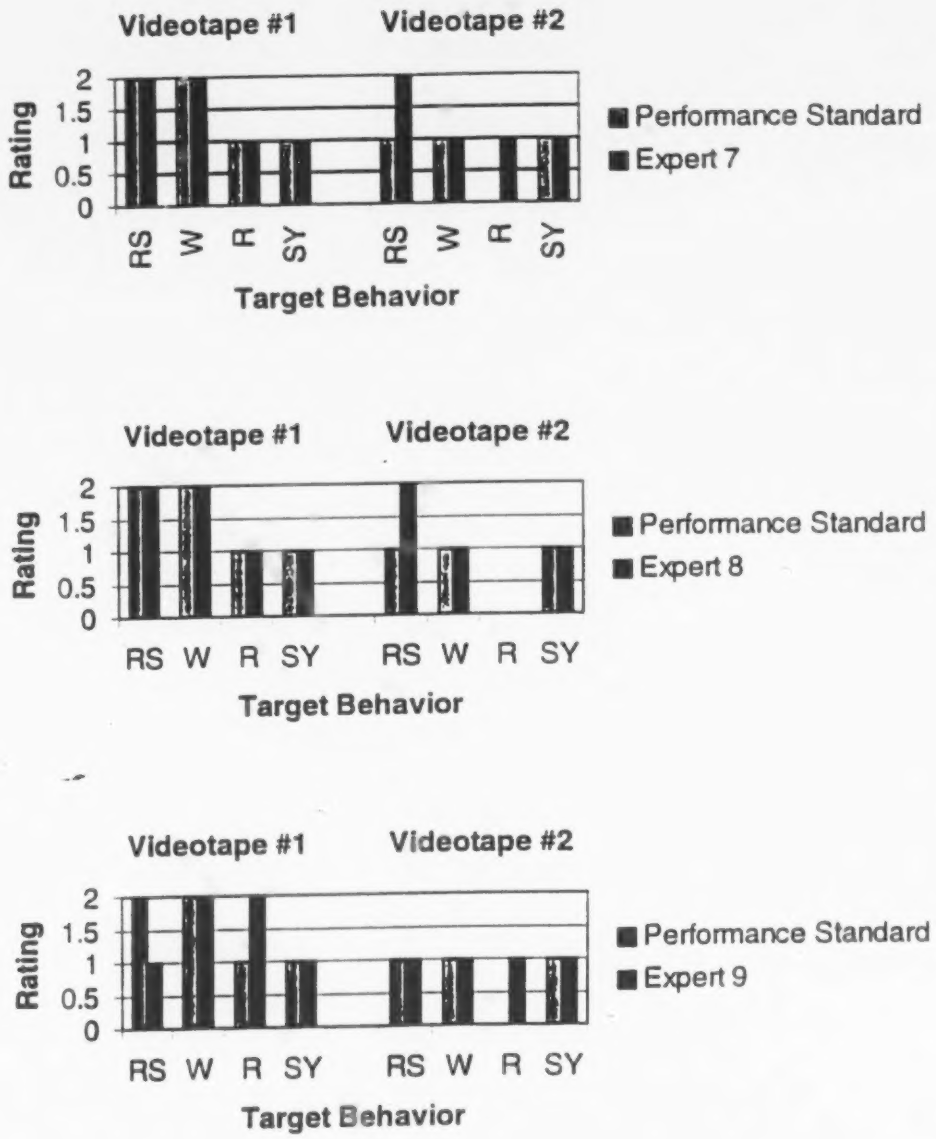


Figure 6 (Cont.). Individual OCD expert's OCD-RS ratings compared to the performance standard for videotapes #1 and #2.

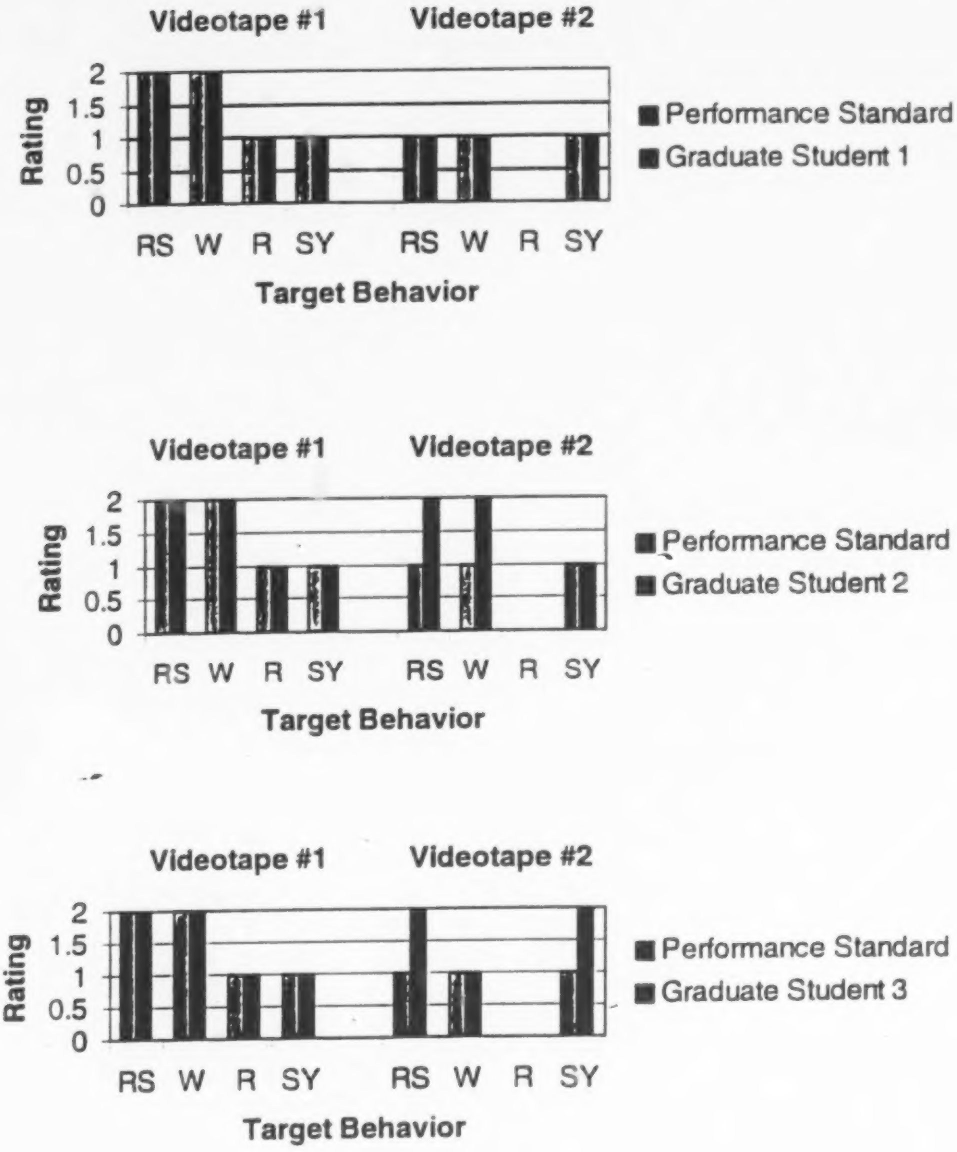


Figure 7. Individual graduate student's OCD-RS ratings compared to the performance standard for videotapes #1 and #2.

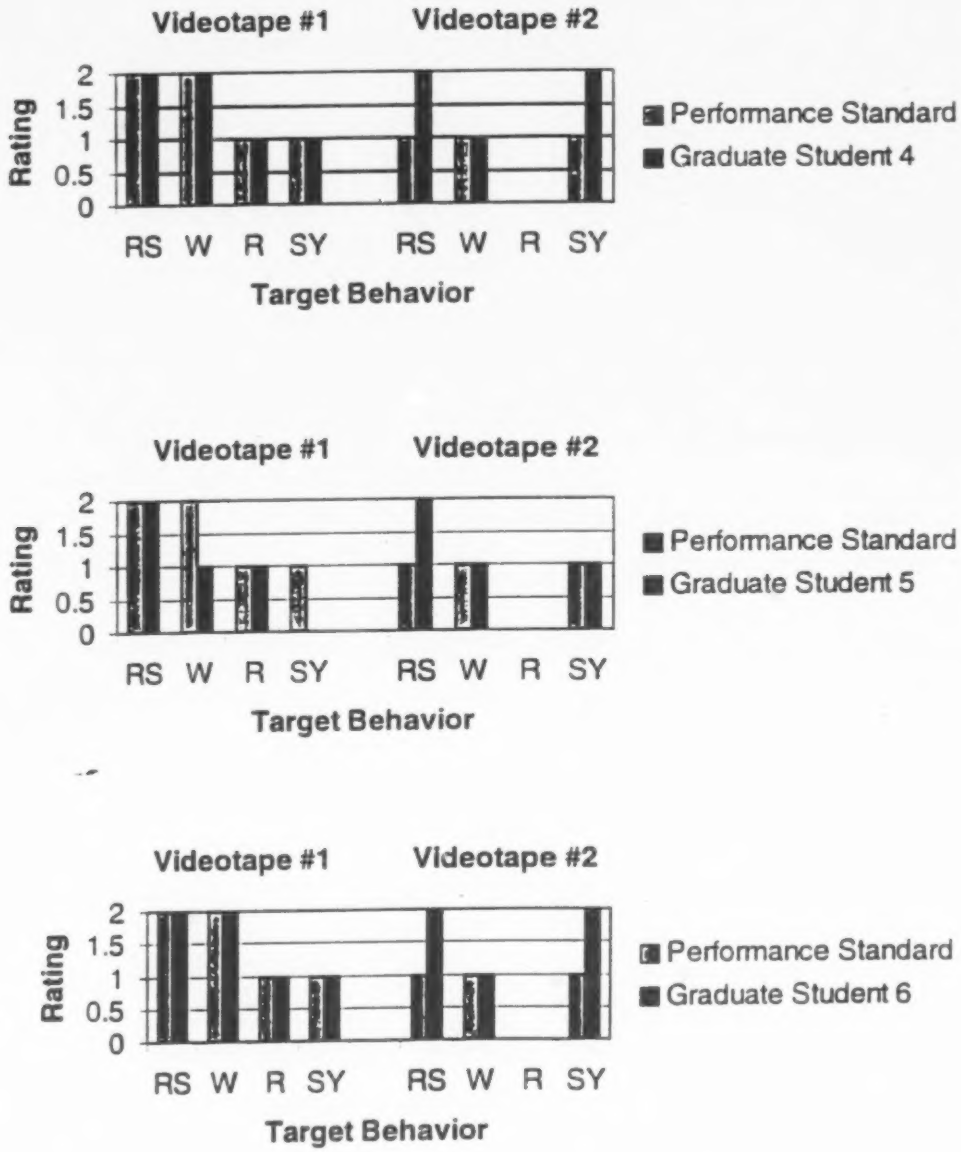


Figure 7 (Cont.). Individual graduate student's OCD-RS ratings compared to the performance standard for videotapes #1 and #2.

variance from PS) the less accurate the collective observations. The RMS, frequency with which accurate ratings occurred, and percent agreement within each target behavior (i.e., item) was calculated for the OCD-RS.

Direct observations. On videotape #1, none of the OCD experts were 100% accurate with respect to the target behaviors with the exception of two experts on Writing and three experts on repeating. However, the variance around the PS was minimal considering the nature of direct observations (i.e., length of tape and intervals). On videotape #1, the range of the RMS was 2.47 to 6.56. In other words, the OCD experts, on average, observed ±2.47 occurrences of writing rituals than the PS; ±3.50 occurrences of symmetry than the PS; ±4.47 occurrences of repeating rituals than the PS; and ±6.56 occurrences of reassurance seeking than the PS. The most accurately observed target behavior for videotape #1 was Writing.

Accuracy improved somewhat on videotape #2 as indicated by the smaller RMS across target behaviors. Additionally, three out of the nine experts were 100% accurate for Writing; five out of the nine were 100% accurate for Repeating; and two out of the nine were 100% accurate for Symmetry. The RMS statistic ranged from 2.12 to 5.07. The OCD experts, on average, observed ±2.12 instances of symmetry than the PS, ±3.86 occurrences of repeating than the PS, ±3.71 occurrences of writing rituals than the PS, and ±5.07 instances of reassurance seeking than the PS.

Symmetry was the most accurately observed target behavior on videotape #2. The descriptive statistics for expert observations using the OCD-DOS on videotapes #1 and #2 are presented in Table 3.

Ratings. The RMS on videotape #1 ranged from .35 (Writing, Repeating, Symmetry) to .50 (Reassurance Seeking). Additionally, 7 out of the 9 experts (78%) rated Reassurance Seeking the same as the PS. In other words, 78% of the OCD experts were 100% accurate for Reassurance Seeking. For Writing, Repeating, and Symmetry, 8 out of the 9 OCD experts, 89% of all OCD experts, were 100% accurate.

For videotape #2, the RMS ranged from .35 to .79. Writing rituals and Symmetry both had an RMS of .35, Repeating had an RMS of .71, and Reassurance Seeking had an RMS of .79. For Reassurance seeking, 4 out of the 9 OCD experts, or 44%, were 100% accurate when compared to the PS. For Writing rituals and Symmetry, 8 out of the 9 experts, or 89%, were 100% accurate when compared to the PS. For the target behavior, Repeating, 5 out of the 9 OCD experts, or 56%, were in exact agreement with the PS. The descriptive statistics for the OCD expert ratings using the OCD-RS on videotapes 1 and 2 may be found in Table 4.

The accuracy of OCD experts for the total scale (i.e., across all target behaviors) was derived by calculating percent agreement with the PS. Overall, the OCD experts were fairly accurate when it came to rating the four

Table 3

Descriptive Statistics by Target Behavior for OCD Expert Observations Using the OCD-DOS on Videotapes 1 and 2.

| | RS | | W | | R | | SY | |
|---|---|---|---|---|---|---|---|---|
| | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 |
| PS | 15 | 9 | 16 | 3 | 3 | 0 | 4 | 7 |
| Mean | 18.11 | 9.44 | 14.78 | 5.22 | 6.11 | 2.33 | 5.56 | 7.44 |
| SE | 1.89 | 1.68 | .70 | .95 | 1.01 | .98 | 1.02 | .69 |
| Median | 20.00 | 10.00 | 15.00 | 4.00 | 6.00 | 0.00 | 5.00 | 8.00 |
| Mode | 20.00 | 10.00 | 15.00[a] | 3.00 | 3.00[a] | 0.00 | 2.00[a] | 7.00[a] |
| SD | 5.67 | 5.05 | 2.11 | 2.86 | 3.02 | 2.96 | 3.09 | 2.07 |
| Range | 7 -25 | 0-15 | 11-17 | 3-11 | 3-12 | 0-7 | 2-11 | 3-10 |
| RMS | 6.56 | 5.07 | 2.47 | 3.71 | 4.47 | 3.86 | 3.5 | 2.12 |

Note. RS= Reassurance Seeking. W=Writing. R=Repeating. SY=Symmetry. RS, W, R, and SY are the target behaviors. V #1= Videotape #1. V #2= Videotape #2. PS = UW-Madison Performance Standard data. SE=Standard Error. SD=Standard Deviation. RMS= Root Mean Square.

[a] Multiple modes exist. Smallest value shown.

Table 4

Descriptive Statistics by Target Behavior for OCD Expert Ratings Using the OCD-RS on Videotapes 1 and 2.

| | RS | | W | | R | | SY | |
|---|---|---|---|---|---|---|---|---|
| | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 |
| PS | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 1 |
| Mean | 1.78 | 1.33 | 1.89 | 1.11 | 1.11 | .44 | 1.11 | 1.11 |
| SE | .15 | .24 | .11 | .11 | .11 | .18 | .11 | .11 |
| Median | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 1 |
| Mode | 2 | 1[a] | 2 | 1 | 1 | 0 | 1 | 1 |
| SD | .44 | .71 | .33 | .33 | .33 | .53 | .33 | .33 |
| Range | 1-2 | 0-2 | 1-2 | 1-2 | 1 | 0-1 | 1-2 | 1-2 |
| RMS | .50 | .79 | .35 | .35 | .35 | .71 | .35 | .35 |
| FREQ | 7/9 | 4/9 | 8/9 | 8/9 | 8/9 | 5/9 | 8/9 | 8/9 |

Note. RS= Reassurance Seeking. W=Writing. R=Repeating. SY=Symmetry. RS, W, R, and SY are the target behaviors. V #1= Videotape #1. V #2= Videotape #2. PS = UW-Madison Performance Standard data. SE=Standard Error. SD=Standard Deviation (Reliability index). RMS= Root Mean Square (Accuracy index). FREQ= the number of experts who agreed with the PS.

[a] Multiple modes exist. Smallest value shown.

target behaviors on the OCD-RS. For Videotape #1, the percent agreement with the PS ranged from 50% to 100% (M = 86%). Six out of the nine experts were 100% accurate when compared to the PS.

On Videotape #2, the OCD experts' accuracy declined. The percent agreement ranged from 25% to 100% (M = 69%). Only two out of the nine OCD experts demonstrated 100% accuracy. Percent agreement statistics relevant to this prediction are displayed in Tables 5 and 6.

Prediction 2: Supported

It was predicted that trained graduate student raters would have moderately high ($r$ = .50 to .80) inter-rater agreement on the OCD-RS. Evidence to support the prediction was gathered by examining rater percent agreement within items and across total scale. Thus, the following questions were addressed: (1) What was the percent agreement among raters for a particular target behavior? (2) What target behavior had the most reliability? and (3) How reliable were pairs of raters across the total scale? Percent agreement within each item across raters was calculated by the number of agreements divided by the total number of raters and multiplied by 100. Item reliability across raters also was assessed by computing the standard deviation. Agreement for pairs of raters was calculated across the total 4-item scale by using percent agreement.

Table 5

Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #1

| OCD Expert | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | PS |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | - | | | | | | | | | |
| 2 | 25 | - | | | | | | | | |
| 3 | 75 | 50 | - | | | | | | | |
| 4 | 75 | 50 | 100 | - | | | | | | |
| 5 | 75 | 50 | 100 | 100 | - | | | | | |
| 6 | 75 | 50 | 100 | 100 | 100 | - | | | | |
| 7 | 75 | 50 | 100 | 100 | 100 | 100 | - | | | |
| 8 | 75 | 50 | 100 | 100 | 100 | 100 | 100 | - | | |
| 9 | 25 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | - | |
| PS | 75 | 50 | 100 | 100 | 100 | 100 | 100 | 100 | 50 | - |

Note. PS=Performance Standard. Percent Agreement was calculated by the number of agreements divided by the total (4) and multiplied by 100.

Table 6

Percent Agreement Between Pairs of OCD Experts' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #2

| OCD Expert | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | PS |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | - | | | | | | | | | |
| 2 | 75 | - | | | | | | | | |
| 3 | 50 | 25 | - | | | | | | | |
| 4 | 100 | 75 | 50 | - | | | | | | |
| 5 | 75 | 50 | 50 | 75 | - | | | | | |
| 6 | 25 | 50 | 25 | 25 | 25 | - | | | | |
| 7 | 50 | 75 | 50 | 50 | 50 | 75 | - | | | |
| 8 | 75 | 50 | 75 | 75 | 75 | 50 | 75 | - | | |
| 9 | 75 | 100 | 25 | 75 | 50 | 50 | 75 | 50 | - | |
| PS | 100 | 75 | 50 | 100 | 75 | 25 | 50 | 75 | 75 | - |

Note. PS=Performance Standard. Percent Agreement was calculated by the number of agreements divided by the total (4) and multiplied by 100.

For Videotape #1, percent agreement across raters ranged from 83% for Writing and Symmetry to 100% for Reassurance Seeking and Repeating. When examining the standard deviations of each item, Repeating and Reassurance Seeking (SD=0) were found to be the most reliable items on the 4-item scale.

On videotape #2, the percent agreement across raters for each item was 50% for Symmetry, 83% for Writing and Reassurance Seeking, and 100% for Repeating. With a standard deviation of 0, Repeating was found to be the most reliable items out of the 4-item scale. Statistical results for this prediction can be found in Table 7.

Percent agreement was computed between pairs of raters to assess inter-rater reliability across the total scale. On videotape #1, percent agreement between pairs of raters (N=15) ranged from 50% to 100% (M % = 83%). Five pairs of raters agreed on 2 out of the 4 items (50%) and 10 pairs of raters agreed on 4 out of the 4 items (100%).

On videotape #2, percent agreement between pairs of raters (N=15) also ranged from 50% to 100% (M % = 68%). Seven pairs of raters agreed on 2 out of the 4 target behaviors (items) (50%); five pairs of raters agreed on 3 out of the 4 (75%) target behaviors; and 3 pairs of raters agreed on 4 out of the 4 target behaviors (100%). Percent agreement for pairs of

Table 7

Descriptive Statistics by Target Behavior across Graduate Student Raters Using the OCD-RS on Videotapes 1 and 2.

| | RS | | W | | R | | SY | |
|---|---|---|---|---|---|---|---|---|
| | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 |
| PS | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 1 |
| Mean | 2 | 1.83 | 1.83 | 1.17 | 1 | 0 | .83 | 1.50 |
| SEM | 0 | .17 | .17 | .17 | 0 | 0 | .17 | .22 |
| Median | 2 | 2 | 2 | 1 | 1 | - | 1 | 1.50 |
| Mode | 2 | 2 | 2 | 1 | 1 | 0 | 1 | 1[a] |
| SD | 0 | .41 | .41 | .41 | 0 | 0 | .41 | .55 |
| Range | 2 | 2 | 1-2 | 1-2 | 1 | 0 | 0-1 | 1-2 |
| RMS | 0 | .91 | .44 | .44 | 0 | 0 | .44 | .77 |
| % AG | 100 | 83 | 83 | 83 | 100 | 100 | 83 | 50 |
| Reliability FREQ | 6/6 | 1/6 | 5/6 | 5/6 | 6/6 | 6/6 | 5/6 | 3/6 |

*Note.* RS= Reassurance Seeking. W=Writing. R=Repeating. SY=Symmetry. RS, W, R, and SY are the target behaviors. V #1= Videotape #1. V #2= Videotape #2. PS = UW-Madison Performance Standard data. SE=Standard Error of Measure. SD=Standard Deviation (item reliability index). RMS= Root Mean Square (Accuracy index). FREQ= the number of experts who agreed with the PS. % AG= percent agreement.

[a] Multiple modes exist. The smallest value is shown.

observers across total scale for videotapes 1 and 2 can be found in Tables 8 and 9, respectively.

Prediction 3: Supported

It was predicted that graduate student observers would have high inter-observer (r = .60 to .90) reliability when comparing their observations on the OCD-DOS. Item reliability across raters was assessed by computing the standard deviation. All other indices of reliability for observers are not suitable due to the continuous nature of direct observations. Like the OCD-RS, Repeating was found to be the most reliable item on the four item scale for both videotapes.

On videotape #1, the range of the standard deviation was .82 to 4.13. In other words, the trained graduate student observers, on average, observed ±.82 occurrences of Repeating than the mean; ±1.83 occurrences of Symmetry than the mean; ±4.13 occurrences of Writing rituals than the mean; and ±2.43 occurrences of Reassurance Seeking than the mean.

On videotape #2, the standard deviation statistic ranged from 0 to 2.99. The graduate student observers were 100% reliable on their observations of Repeating. The graduate student observers, on average, observed ±1.51 instances of Writing rituals than the mean; ±1.35 occurrences of Symmetry than the mean; and ±2.99 occurrences of writing rituals than the mean. The descriptive statistics for graduate student

Table 8

Percent Agreement Between Pairs of Graduate Students' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #1

| Rater | 1 | 2 | 3 | 4 | 5 | 6 | PS |
|---|---|---|---|---|---|---|---|
| 1 | - | | | | | | |
| 2 | 100 | - | | | | | |
| 3 | 100 | 100 | - | | | | |
| 4 | 100 | 100 | 100 | - | | | |
| 5 | 50 | 50 | 50 | 50 | - | | |
| 6 | 100 | 100 | 100 | 100 | 50 | - | |
| PS | 100 | 100 | 100 | 100 | 50 | 100 | - |

Note. Group 1= Raters 1 -3. Group 2= Raters 4 - 6. PS=Performance Standard. Percent Agreement was calculated by the number of agreements divided by the total (4) and multiplied by 100.

Table 9

Percent Agreement Between Pairs of Graduate Students' OCD-RS Ratings and the Performance Standard's OCD-RS Ratings on Videotape #2

| Rater | 1 | 2 | 3 | 4 | 5 | 6 | PS |
|---|---|---|---|---|---|---|---|
| 1 | - | | | | | | |
| 2 | 50 | - | | | | | |
| 3 | 50 | 50 | - | | | | |
| 4 | 50 | 50 | 100 | - | | | |
| 5 | 75 | 75 | 75 | 75 | - | | |
| 6 | 50 | 50 | 100 | 100 | 75 | - | |
| PS | 100 | 50 | 50 | 50 | 75 | 50 | - |

Note. Group 1= Raters 1 -3. Group 2= Raters 4 - 6. PS=Performance Standard. Percent Agreement was calculated by the number of agreements divided by the total (4) and multiplied by 100.

observations using the OCD-DOS on videotapes #1 and #2 is presented in Table 10. Item variance around the mean, or standard deviation, is presented in the row labeled SD.

Prediction 4: Supported

It was predicted that graduate students' observations and ratings would be highly accurate when compared to the performance standard's observations and ratings.

Direct Observations. Evidence was found to support this prediction by calculating the RMS for each target behavior (item). Accuracy across the total scale was calculated by using Pearson correlations.

On videotape #1, the range of the RMS across the four target behaviors was .89 to 7.69. In other words, the trained graduate student observers, on average, observed ±.89 occurrences of Repeating than the PS; ±1.84 occurrences of Symmetry than the PS, ±7.16 occurrences of Writing rituals than the PS; and ±7.69 occurrences of Reassurance Seeking than the PS.

On videotape #2, the RMS statistic ranged from 0 to 6.40. The trained graduate student observers, on average, observed ±1.67 instances of Writing rituals than the PS; ±3.22 occurrences of Symmetry than the PS, and ±6.40 occurrences of Reassurance Seeking than the PS. The trained graduate observers were 100% accurate in their observations of Repeating. In sum,

Table 10

*Descriptive Statistics by Target Behavior across Graduate Student Observers Using the OCD-DOS on Videotapes 1 and 2.*

| | RS | | W | | R | | SY | |
|---|---|---|---|---|---|---|---|---|
| | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 | V #1 | V #2 |
| PS | 15 | 9 | 16 | 3 | 3 | 0 | 4 | 7 |
| Mean | 21.67 | 14.17 | 14.33 | 2.33 | 2.67 | 0 | 3.83 | 9.67 |
| SE | .99 | 1.22 | 1.69 | .61 | .33 | 0 | .75 | .56 |
| Median | 21.50 | 14.50 | 15.50 | 3.00 | 2.50 | - | 3.50 | 10.00 |
| Mode | 22.00 | 16.00 | 18.00 | 3.00 | 2.00 | 0 | 2.00[a] | 8.00[a] |
| SD | 2.42 | 2.99 | 4.13 | 1.51 | .82 | 0 | 1.83 | 1.36 |
| Range | 19-26 | 10-18 | 8-18 | 0-4 | 2-4 | 0 | 2-6 | 8-11 |
| RMS | 7.69 | 6.40 | 7.16 | 1.67 | .89 | 0 | 1.84 | 3.22 |

*Note.* RS= Reassurance Seeking. W=Writing. R=Repeating. SY=Symmetry. RS, W, R, and SY are the target behaviors. V #1= Videotape #1. V #2= Videotape #2. PS = UW-Madison Performance Standard data. SE=Standard Error. SD=Standard Deviation (Reliability index). RMS= Root Mean Square (Accuracy index).

[a] Muliple modes exist. Smallest value is shown.

Repeating was observed with the most accuracy for videotapes #1 and #2. The descriptive statistics for graduate student observations using the OCD-DOS on videotapes #1 and #2 is presented in Table 10.

Accuracy across the total 4-item OCD-DOS was assessed by calculating Pearson correlations between the PS and the graduate students. For videotape #1, correlations ranged from .78 to .98 (*M* = 91). Slightly higher correlations were discovered for videotape #2. Pearson correlations ranged from .95 to .99 (*M* = 97). These correlations are presented in Table 11.

*Ratings.* Evidence was found to support the prediction that trained graduate students would be highly accurate when compared to the PS. Evidence was found to support this prediction by calculating the RMS, frequency ("hit rate") with which accurate ratings occurred, and percent agreement within each target behavior (i.e., item).

For videotape #1, the RMS ranged from 0 (Repeating and Reassurance Seeking) to .44 (Writing and Symmetry). All of the six trained graduate raters were 100% accurate when compared to the PS for Repeating and Reassurance Seeking. Additionally, 5 out of the 6 trained graduate student raters rated Writing rituals and Symmetry with 100% accuracy.

For videotape #2, the RMS was 0 for Repeating, .44 for Writing, .77 for Symmetry, and .91 for Reassurance Seeking. For Reassurance seeking,

Table 11

*Pearson Correlations between the Performance Standard's and the Graduate Student's Observations on the OCD-DOS on Videotapes 1 and 2.*

Videotape 1

Observers

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| PS | .98* | .91 | .98* | .91 | .78 | .89 |

Videotape 2

Observers

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| PS | .99** | .96* | .99** | .96* | .95* | .99* |

*Note*. Group 1= Observers 1-3; Group 2 = Observers 4-6. PS= Performance Standard.

* *p* < .05

** *p*< .01

1 out of the 6 graduate students was in exact agreement with the PS. For Writing rituals, 5 out of the 6 graduate students were in 100% agreement with the PS. For Repeating, all 6 raters were in agreement with the PS. For Symmetry, 3 out of the 6 raters were in agreement with the PS. The descriptive statistics for the graduate student ratings using the OCD-RS on videotapes #1 and #2 may be found in Table 7.

The accuracy of graduate student raters for the total scale (i.e., across all target behaviors) was derived by calculating percent agreement between an individual rater and the PS. For Videotape #1, the percent agreement with the PS ranged from 50% to 100%. Five out of the six graduate student raters had 100% agreement with the PS. That is, 5 out of the 6 graduate students were 100% accurate when compared to the PS. The remaining graduate student rater was 50% accurate (2 out of 4 items correct). On Videotape #2, one out of the six graduate student raters was 100% accurate when compared to the PS. Another rater was found to be 75% accurate when compared to the PS. The remaining four raters, were 50% accurate when compared to the PS. Percent agreement statistics relevant to this hypothesis are displayed in Tables 8 and 9.

Accuracy also was estimated across the total scale by computing Pearson correlations between the PS and each individual graduate-student rater. For videotape #1, correlations ranged from .71 to 1.00 (M = .95;

Mode = 1.00). For videotape #2, correlations ranged from .82 to 1.00 (M = .73, Mode = .87). Pearson correlations between graduate student raters and the PS can be found in Table 12.

Prediction 5: Supported

Evidence was found to support the prediction that an observer's observation and his/her own rating would achieve or exceed moderate correlation. Pearson correlations were calculated to demonstrate the strength of the relationship between observations and ratings.

Graduate students. For videotape #1, the correlations between a graduate student's OCD-DOS observation and his/her own OCD-RS ratings ranged from .87 to .98, with a mean of .94. For videotape #2, the correlations ranged from .66 to .97, with a mean of .83. These correlational results can be found in Table 13.

OCD Experts. For videotape #1, the correlations between an OCD expert's OCD-DOS observation and his/her own OCD-RS ratings ranged from .70 to .99, with a mean of .91. For videotape #2, the correlations ranged from .66 to .98, with a mean of .86. A Pearson correlation for expert 2 could not be calculated for videotape #1, and a Pearson correlation could not be computed for expert 2 and expert 9 on videotape #2 because there was no variability among their ratings. These correlational results can be found in Table 14.

Table 12

Pearson Correlations between the Performance Standard's and the Graduate Student's Ratings on the OCD-RS on Videotapes 1 and 2.

Videotape 1

Raters

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| PS | 1.00** | 1.00** | 1.00** | 1.00** | .71 | 1.00** |

Videotape 2

Raters

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| PS | 1.00** | .87 | .87 | .87 | .82 | .87 |

Note. Group 1= Raters 1-3. Group 2 = Raters 4-6. PS= Performance Standard.

* $p < .05$

** $p < .01$

Table 13

Pearson Correlations between a Graduate Student's OCD-DOS Observation and His/Her Own OCD-RS Rating on Videotapes 1 and 2.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | Videotape 1 | | | | |
| Observer | 1 | 2 | 3 | 4 | 5 | 6 | PS |
| | 98** | 93 | .98* | .94 | .87 | .92 | .99** |
| | | | Videotape 2 | | | | |
| Observer | 1 | 2 | 3 | 4 | 5 | 6 | PS |
| | .77 | .66 | .97* | .87 | .79 | .97 | .79 |

*Note*. Group 1= Raters 1-3. Group 2 = Raters 4-6. PS= Performance Standard.

* *p* < 05

** *p*< 01

Table 14

Pearson Correlations between an OCD Expert's OCD-DOS Observation and His/Her Own OCD-RS Rating on Videotapes 1 and 2.

Videotape 1

| Observer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | PS |
|---|---|---|---|---|---|---|---|---|---|---|
| | .70 | $-^{a}$ | .87 | .99* | .98* | .86 | .96* | .93 | .99* | .99* |

Videotape 2

| Observer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | PS |
|---|---|---|---|---|---|---|---|---|---|---|
| | .76 | $-^{a}$ | .88 | .66 | .88* | .95 | .93 | .98* | $-^{a}$ | .79 |

*Note*. Group 1= Raters 1-3. Group 2 = Raters 4-6. PS= Performance Standard.

$-^{a}$ cannot be computed because one of the variables was constant.

* $p < .05$

** $p < .01$

Prediction 6: Supported

It was predicted that the new measures, the OCD-RS and OCD-DOS, would be sensitive to change. This prediction was evaluated by calculating an effect size (ES). The effect size was calculated by dividing the difference in phase means (videotape #1 - videotape #2) by the baseline standard deviation (videotape #1) (Busk & Serlin, 1992). This approach places no assumptions on the data regarding the distributional form or homogeneity of the variance. It was predicted that there would be a decrease in behavior from videotape #1 to videotape #2. This decrease in behavior should have been observed in the following target behaviors: Reassurance Seeking, Writing, and Repeating. The remaining target behavior, Symmetry, increased slightly from videotape #1 to videotape #2 as evidenced in observations(i.e., Videotape #1= 4 occurrences; Videotape #2=7 occurrences) but not in ratings (i.e., both videotapes had a PS rating of 1).

Direct observations. For OCD experts, the ES ranged from -.61 to 4.53. Collectively, the OCD experts' OCD-DOS observations from videotape #1 to videotape #2 ranged from a moderate effect size (SY= -.61) to a large effect size (RS, W, R). Across graduate student observers, the OCD-DOS was found to be sensitive to change for all target behaviors in the expected directions. The effect sizes ranged from -3.19 to 3.66. The binomial sign test (Marascuilo & Serlin, 1988) was applied to the raw data to detect statistical significance. The difference between videotape #1 and #2 was statistically significant for RS, W, R across OCD experts and statistically significant for RS, W, R, and SY across graduate students ($\alpha = .05$). The effect sizes by target behavior may be found in Table 15.

Ratings. For OCD experts, the ES ranged from 0 (SY) to -2.36. Collectively, the OCD experts' OCD-RS rating from videotape #1 to videotape #2 was not sensitive to moderately sensitive to change as evidenced by the ES. Across graduate student raters, the OCD-RS was found to be moderately sensitive to change as evidenced by the moderate effect sizes. The effect sizes ranged from 1.63 to -1.63. ES for ratings should be interpreted with caution due to the forced choice nature of the rating format as well as the small number of choices (N=3). See Tables 13 for the OCD experts' and graduate students' results on the OCD-RS.

The binomial sign test also was applied to the raw rating scale data to see if the scores reported were statistically significant. Across OCD experts, the only statistically significant difference between videotape #1 and #2 was Writing ($\alpha = .05$). Across graduate students the only statistically significant difference that occurred from videotape #1 to videotape #2 was for Repeating ($\alpha = .05$).

Table 15

Effect Sizes for OCD Experts and Graduate Students by Target Behavior on the OCD-DOS and OCD-RS

| | OCD-DOS | | | |
|---|---|---|---|---|
| | RS | W | R | SY |
| OCD Experts | 1.53 | 4.53 | 1.25 | -.61 |
| Students | 3.10 | 2.91 | 3.66 | -3.19 |
| | **OCD-RS** | | | |
| | RS | W | R | SY |
| OCD Experts | 1.02 | 2.36 | 2.03 | 0.00 |
| Students | .17 | 1.63 | 1.00 | -1.63 |

Note. RS= Reassurance Seeking. W= Writing. R=Repeating SY=Symmetry

# Chapter 5

# Discussion

The purpose of this study was to develop a rating scale and direct observation measure that could be used in the home or school setting to monitor the treatment of childhood onset OCD. The assessment of OCD has focused primarily on indirect methods of assessment such as rating scales or checklists; however, current OCD rating scales have several limitations such as a lack of adequate standardization as well as limited reliability and validity (i.e., content validity and accuracy). In addition, there are are no systematic direct observation measures for the outcome assessment of OCD despite repeated requests for this type of assessment methodology (Como, 1997; Francis & Gragg, 1996; Taylor, 1995). Given that there is a need for the development of accurate, symptom specific outcome assessment measures, this study served to extend the current OCD literature base by developing two new OCD outcome assessment measures.

An alternative scale validation model was used to validate the much needed quality OCD outcome assessment measures developed in this study. This alternative scale validation model has been referred to in the literature as the accuracy-reliability framework or paradigm, and was originally conceptualized by Cone (1981; 1992) and then extended by Robertson

(1993) and Racine (1994) in their work with children's social behavior. Seymour and Kratochwill (1998) applied this framework to the childhood disorder of selective mutism. In this study, the performance standard, or incontrovertible index, was established via scripted, analog videotapes and consensual expert agreement on the rating scale (OCD-RS) and the direct observation system (OCD-DOS). Extreme care was taken in this study to obtain a performance standard from which an accuracy measure could be meaningfully derived. Data from the OCD field experts and the trained graduate students were then compared to this performance standard. High agreement with the performance standard and among participants would suggest that the OCD-RS and the OCD-DOS are both accurate and reliable.

## Generalizations and Implications of the Findings

Several findings of this study are noteworthy with respect to ratings and direct observations in the areas of reliability, accuracy, relationship, and sensitivity. An examination of these findings and their implications for researchers and practitioners is discussed in the remainder of the chapter.

### Reliability of Ratings and Direct Observations

Ratings. The moderate to moderately high percent agreements between graduate students' OCD-RS ratings indicated that there was relative agreement of content for both videotapes with regard to the frequency of the four target behaviors sampled by the OCD-RS. Overall, graduate students were more reliable than the OCD experts. Agreement between pairs of graduate student raters across the total scale were moderately high for both videotapes (videotape #1: M % = 83%, and videotape #2: M % = 68%). Experts displayed moderate to moderately high percent agreement across the total scale (videotape #1: M % = 74%, and videotape #2: M % = 58%). These results can be considered impressive given that exact percent agreement was used and that such a small sample of items comprised the total scale. However, past researchers do not report inter-rater reliability in percent agreement but rather Pearson correlations. Thus, the results from this study are not readily interpretable in the context of previous literature.

In terms of the scale itself, Repeating and Reassurance Seeking were found to be the most reliable items for videotape #1 among graduate student raters, and Repeating was found to be the most reliable item on videotape #2. Among OCD experts, Writing, Repeating, and Symmetry were found to be equally reliable on videotape #1; and Writing and Symmetry were found to be the most reliable items on videotape #2. This difference in reliable items between graduate students and OCD experts may be due to the varying degrees of familiarity and conceptualization of OCD. That is, graduate students might have found it easier to discriminate between Repeating and Reassurance Seeking as compared to Writing and Symmetry; whereas, OCD experts might have been more in tune to the topography of all behaviors as seen in research and practice.

Direct observations. Another area of interest in this study involved inter-observer reliability. Unlike the OCD-RS ratings, exact percent agreement was not used for direct observations due to the continuous nature of observational data. However, if one examines the observation data for occurrence/non-occurrence scores which is the type of percent agreement previously reported in the literature, graduate students were in 100% agreement for videotape #1. On videotape #2, five out of the six graduate students had100% agreement. The experts demonstrated 100% agreement on videotape #1, but their performance declined on videotape #2. On videotape #2, four out of the nine expert observers indicated that they saw Repeating rituals when this behavior did not occur on the videotape. For the remaining three target behaviors, however, the OCD experts demonstrated 100% agreement.

The results for trained graduate student observers were consistent with the past research, if not higher. For example, Racine (1994) documented a mean percent agreement between pairs of observers of 76% and 84% for her videotapes; and Robertson (1993) demonstrated a mean percent agreement between pairs of observers of 75%. Therefore, if given optimal observation situations, reliable results can be produced for both ratings and observations.

Summary and implications. The higher performance by the trained graduate students on the OCD-RS and the OCD-DOS may have been due to the criterion mastery training approach that was used in the current study. Specifically, all graduate students participated in a training program that incorporated direct instruction and a comprehensive manual. In addition, graduate students had to attain 85% accuracy on the OCD-DOS while observing videotapes that featured other target children and adolescents with OCD. It appears that this approach may have fostered better attendance to the target behaviors; and in turn, yielded higher results when compared to the experts who had not been similarly trained. While experts were given the same comprehensive manual, they were left to their own discretion as to its use before completing the OCD-DOS and OCD-RS. The question remains as to how they read and studied the manual before they completed the required tasks.

The assessment of OCD requires considerable skill because implicit in the diagnosis of OCD is the presence of an affective or cognitive component (e.g., fear of contamination or need for exactness) that drives the compulsions. With this in mind, it is understandable that four of the "untrained" experts seemed to confuse the target behaviors of Repeating and

Symmetry as the repetition of an action occurs when a student is trying to arrange or order something in a symmetrical fashion. Thus, in addition to observing a behavior, the observer must make some interpretation about the behavior. Interestingly, the "OCD experts" proved they were fallible under an artificially constructed situation which may or may not approximate natural conditions. This finding has implications for practice in that, often clinicians may err in making attributions about behavior. It is important to remember that this accuracy-reliability paradigm is not immune from the same problems often encountered with the use of traditionally validated scales. Overall, training seems to reduce errors in attribution, and perhaps, "experts" should not be excluded from this type of training in the future. In other words, it seems that being recognized as an "expert" is not necessarily as good as being a "trained" expert.

## Accuracy of Ratings and Direct Observations

For behavioral measures, such as the OCD-RS and OCD-DOS, Foster and Cone (1995), contend that two types of representational validity need to be established: content validity and accuracy. Content validity examines the extent to which a measure operationally defines the behavior/response class being assessed. Also, the obtained scores on a particular measure of behavior should reflect actual behavior. Content validity was established in this study by operationally defining four target behaviors based on current

OCD research and direct observations of children and adolescents who had been diagnosed with OCD. Additionally, these four behaviors were scrutinized by an OCD expert to brainstorm all possible examples and non-examples of the target behaviors for the purpose of constructing a performance standard, or incontrovertible index. Scores on the OCD-RS and OCD-DOS reflect the frequency of the target behaviors. The higher the number, the more frequently the behavior occurs.

The main focal point of this study was the concept of accuracy. Not only was it important to establish accuracy from a theoretical standpoint but also from a practical standpoint. One current evaluation of OCD assessment instruments cited the need for more accurate assessment of OCD symptoms (Como, 1997). The construct of accuracy examines the instrument's ability to reflect "true behavior" as defined by the performance standard. For the most part, accuracy has been ignored by OCD assessors because it is often difficult to obtain an incontrovertible index, or performance standard. Hence, there are no OCD rating scales or direct observation measures that have established accuracy.

Cone (1992) stipulated that to establish the accuracy of a behavioral measure, the researcher must provide clear, explicit directions and procedures for using the instrument, and a performance standard so as to compare outcome data. In this study, clear, concise, and explicit rules were detailed in the observation and rating manual given to all participants. In addition to rating and observation guidelines, this manual consisted of a literature review of OCD and the assessment of OCD. Detailed operational definitions with examples for the four target behaviors and a self-study behavior codes quiz also were included in the manual.

Ratings. OCD experts were fairly accurate when it came to rating the four target behaviors on the OCD-RS. For videotape #1, the experts' accuracy ranged from 50% to 100% (M = 86%). For videotape #2, their accuracy was somewhat less (M = 69%). An analysis of ratings between OCD experts and the performance standard on the OCD-RS revealed that six out of the nine were 100% accurate for videotape #1. For videotape #2, only two experts were 100% accurate. This decline in accuracy could be due to an order effect which was not controlled for as it was with the graduate students. Given that there were higher frequency behaviors in videotape #1, it is possible that the OCD experts might have expected the same frequency of behaviors to occur in videotape #2.

Trained graduate students were highly accurate when compared to the performance standard. In many cases, 100% accuracy was obtained for several target behaviors. The accuracy of the ratings declined somewhat on videotape #2. Across the total scale, moderately high to high Pearson correlations and percent agreements were found (videotape #1: M r = 95 and M % = 92%. and videotape #2: *M* *r* = .88 and *M* % = 63%). These findings are somewhat higher than found in Racine (1994) and Seymour and Kratochwill (1998). These results may reflect the combined emphasis on observation and rating in the training manual.

*Direct observations*. The OCD experts attained a moderate level of accuracy. Overall, however, the variance around the performance standard was minimal considering the nature of direct observations (i.e., length of tape and intervals). In examining occurrence/non-occurrence data, all the experts observed the target behaviors with differing frequencies. In many cases, OCD experts were not that far away from the performance standard. The most accurately observed target behavior for videotape #1 and videotape #2 was Writing (*RMS*= 2.47) and Symmetry (*RMS*=2.12), respectively.

Trained graduate students were more accurate than the OCD experts. Repeating was observed with the most accuracy for videotapes #1 and #2. Moderately high to high correlations were discovered across the total OCD-DOS for Videotape #1 (*M* *r* = .91). Slightly higher correlations were discovered for videotape #2 (*M* *r* = .97).

Correlations between trained graduate students and the performance standard in this investigation were higher than past studies (Robertson, 1993. Racine, 1994). Robertson (1993) and Racine (1994) demonstrated mean correlations of .89 and .74 for the same videotape, respectively.

Racine used an additional videotape and found a mean r of .87. Seymour and Kratochwill (1998) found a mean r of .67. No other empirical research has assessed this dimension of representational validity for direct observations.

Summary and implications. Across graduate students and OCD experts, accuracy declined on the OCD-RS for videotape #2. In actuality, three (RS, W, R) out of the four target behaviors decreased in frequency from videotape #1 to videotape #2. Since the experts were not counterbalanced like the graduate students, this condition could have resulted in observer expectancy from previous knowledge of higher frequency behaviors seen on videotape #1. In other words, the OCD experts could have had false expectations and over-recorded behavioral observations on videotape #2, thus elevating their numerical ratings.

However, this possible order effect does not explain the graduate students' decline in accuracy. Again, skill in discriminating the observed target behavior from a non-target behavior might have been a factor. Graduate students tended to rate the behaviors with higher frequency than actually occurred on videotape #2. To illustrate the skill needed in rating OCD behavior, one needs to examine the operational definitions. For example, Reassurance Seeking is defined by the type of question asked. Thus, graduate students also may have been victim to expectancy bias by assuming all questions asked by the target adolescent fell into the category of Reassurance Seeking.

The observations on the OCD-DOS by trained graduate students were fairly accurate, especially given the stringent nature of these observations. For example, participants in this study were not given the opportunity to stop and rewind the videotapes to enhance their accuracy. In this respect, the observation situation approximated a naturalistic observation period. One would expect the results to be lower given this condition of the study. Again, these findings underscore the importance of criterion-mastery training for observers especially since rules and procedures are linked directly to accuracy (Cone, 1992).

The criterion-mastery approach to training observers used in the Robertson (1993) and Racine (1994) studies highlights the importance of training on both rating scale and direct observation measures. This aspect becomes even more salient when one examines the expert data from the present study. In examining all the predictions, the OCD experts, who participated with a training manual but without training, were less accurate and reliable than the trained graduate students. In essence, the combination of structured training and a manual appeared to elevate the knowledge and awareness of pre-doctoral students to the status of expert. It seems then that expertise in knowledge, reserach, and clinical practice domains as in the case of the experts may not equate with high performance using observations and rating instruments. Furthermore, traditional standards of expertise may not obviate the need for specific training as used with the graduate students. All individuals who have been trained using a criterion mastery approach for the use of scales developed in accordance with the accuracy-reliability paradigm, might be considered the experts in future studies.

One also could speculate that the OCD experts might not have agreed with the operational definitions of the four target behaviors, which could have created observer drift. It also is possible that the decline in accuracy among experts could be attributed to the experts' doctoral/medical training and/or continued professional development choices. For example, those experts who were trained as educational psychologists tended to do better than the clinical psychologists and the psychiatrists. Also, their continued professional development choices might not have focused on observational and rating methodology, whereas, current educational psychology graduate programs place a heavy emphasis on direct observations and ratings in assessment training.

## Relationship between Direct Observations and Ratings

Another major finding in this study concerned the correlation between an observer's OCD-DOS and his/her own OCD-RS rating. Recall that graduate students demonstrated high correlations for videotape #1 (M r = .91) and moderately high correlations for videotape #2 (M r = .83). Similar results were documented for OCD experts (videotape #1: M r = .94; and videotape #2: M r = .86). The results from this study are substantially higher than the weak to moderate correlations reported in previous research studies (Merrell, 1993; Elliott et al., 1988; Reed & Edelbrock, 1983; Robertson, 1993; Racine, 1994; Seymour & Kratochwill, 1998). These results are highly consistent with Cone's (1981) contention that the accuracy of a given instrument can be enhanced by comparing it to another form of measurement. These results corroborate the notion that when two instruments are developed in a concerted, coordinated effort, higher correlations can be achieved.

Summary and implications. In examining the concept of accuracy and the practice of multiple method assessment, this study relied upon an inference made by Cone (1981) that if two assessment measures are shown to be accurate then they are related. Given the moderately high correlations documented in the past, Racine (1994) made a plea for continued use of multiple types assessment devices due to the unaccounted variance between measures. Given the high correlations demonstrated in this study, perhaps practitioners could reconsider using multiple assessment devices if a high level of agreement was established a priori. This strategy would be a time efficient, cost effective way in which to approach school based practice. After all, if a stringent level of agreement is attained between measures, it would be redundant to use two assessment devices per person that could give the same result. Perhaps this tactic would allow the practitioner more time to gather other types of data such as self-report data and teacher interviews, while the other *trained* teacher and/or parent are collecting rating scale data that is similar to observation data. However, until such relationships are documented, practitioners should continue using a multiple method, multiple informant across multiple settings approach to assessment.

## Sensitivity to Change

Given that the purpose of this study was to develop two accurate instruments for treatment monitoring, sensitivity to change was an important component of this investigation as well as a strong test of the instruments. In essence, videotape #1 could be considered the baseline phase and videotape #2 could be considered the treatment phase in a single-case design study. The behaviors reflected on the videotape were scripted to reflect change from baseline to treatment. Reassurance Seeking, Writing, and Repeating were designed to reflect a decrease in behavior from videotape #1 to videotape #2. The frequency of Symmetry increased slightly from videotape #1 to videotape #2. This slight increase should be detected in direct observations, but not in the ratings. Therefore, no significant effect size was expected for Symmetry on the OCD-RS.

Ratings. For the experts, the OCD-RS was sensitive to change for Reassurance Seeking, Writing, and Repeating and not sensitive to change for Symmetry. For graduate students, the OCD-RS did not display sensitivity in the directions expected. Symmetry demonstrated the highest sensitivity when in reality there should have been no difference between videotape #1 to videotape #2. Writing and Repeating were moderately sensitive to change, and Reassurance Seeking demonstrated low sensitivity to change.

Direct observations. Across OCD experts, the OCD-DOS demonstrated moderate sensitivity to change for Symmetry and high sensitivity to change for Reassurance Seeking, Repeating, and Writing. Across graduate student observers, the OCD-DOS was sensitive to change in the expected directions for all four target behaviors.

Summary and implications. In many respects, one would expect observations to be more sensitive to change given the continuous nature of observations as compared to the forced choice nature of ratings. This aspect of perceiving sensitivity to change between videotaped scenarios on a likert type rating scale might be where the years of experience and expertise serve the OCD experts better than graduate students. Perhaps, the OCD-RS sensitivity could be enhanced by creating more forced choice responses with a constricted range of frequency. With more frequency choices on the OCD-RS, it probably would detect more/less change. This issue of documenting sensitivity to change by item is an important contribution to the OCD literature because these data are heretofore non-existent. In terms of treatment, many clinicians and researchers assert that there has been a positive change, but they fail to indicate what symptoms have changed. Therefore, it is important that researchers and clinicians establish sensitivity to change for the assessment instruments they are using. This facet alone could help shed light on whether certain symptoms respond preferentially to certain treatments and perhaps even clarify why some children respond much better to treatment than others.

## Limitations and Suggestions for Future Research

Careful consideration must be given to the generalizations from this study to previous research because of the different statistical techniques which were used. The descriptive statistical analyses used here are not directly comparable to the traditional scale validation studies. However, the present study replicated all of Robertson's (1993), Racine's (1994), and Seymour and Kratochwill's (1998) predictions with respect to trained graduate student observers.

A possible limitation in this study is the use of a scripted videotape to approximate natural stimuli and the diagnosis of OCD. However, one can argue that the results of this study would generalize to naturalistic observations because of the naturalistic qualities of the videotape and the manner in which observations were conducted. The participants in this study were not allowed to stop and rewind the tape. Thus, the observations flowed continuously as in a live observation situation. In a recent, similar investigation, Seymour and Kratochwill (1998), allowed participants to stop and rewind the videotape as many times as possible yet the correlations were lower. This lends credibility to the belief that results based on analog videotape would would generalize to live observations. Nonetheless, this aspect should be tested.

Although the combined (experts and graduate students) sample size was similar to those used by the majority of previous researchers who investigated similar questions, it is small for the purposes of testing generalizability across raters and observers. Another limitation of this study was that experts were not randomly assigned to two groups like the trained graduate students. Since they were not counterbalanced, an order effect could have occurred. The generalizability of the OCD-DOS and OCD-RS are limited in that each individual case of OCD is different; therefore, different target behaviors will have to be developed for each person. In addition, the content validity will have to be re-established for each new OCD case. Ironically, this limitation of generalizability can also be perceived as a strength in that these two instruments are the first individual-specific OCD assessment instruments ever developed for the outcome assessment of OCD.

Since some items demonstrated less accuracy and not enough sensitivity to change, it will be important to extend the current research study by examining the operational definitions. In addition, since there was a decline in accuracy and reliability across experts and graduate students for ratings between videotapes #1 and #2, researchers might need to emphasize and reflect on the importance of remaining objective for each *independent* observation. The addition of more forced choice responses might need to be created for the rating scale since it displayed lower effect sizes than the observations.

Admittedly, this study is just the first step in the development of two accurate and reliable case specific OCD assessment instruments. Given that this study was conducted with actors, the next logical step would be to replicate studies with several adolescents who actually have OCD. This process could involve the school psychologist or mental health clinician conducting conjoint consultation (Sheridan, Kratochwill, & Bergan, 1996) with parents and teachers to facilitate the identification and definition of problematic, overt OCD behaviors. In essence, this phase establishes content validity. Accuracy would be established through criterion training.

Thus, parents and teachers would collect the data across baseline and treatment phases. Once this process was established with several cases, a logical next step would be to extend it to other childhood disorders. This criterion-referenced or accuracy approach to test development could significantly decrease the amount of time and resources expended on sampling hundreds of children, and it also could serve to more directly establish construct validity as well as reliability. See Figure 8 for an outline of developmental sequence for future research.

## Conclusion

This study demonstrated that the relationship between direct observations and ratings in the assessment of observable compulsions can be high when conditions are optimal. The method variance typically associated between indirect and direct assessment measures is more likely to be reduced when care is taken in developing a rating scale and an observation system. This coordination of instrumentation is important in settings, such as school, home, and community, that rely on multiple assessment techniques.

Most importantly, this study illustrated how psychometrically and qualitatively sound outcome assessment measures for a childhood disorder, OCD, can be developed when using an alternative scale validation paradigm. By developing a performance standard, the accuracy of the rating scale and direct observation system could be attained. Clearly, more OCD researchers can provide accuracy information by developing a performance standard. This study contributes to the current OCD assessment domain in that it provides a way to accurately assess specific OCD behaviors. In addition, it provided an example of how to incorporate direct observations into the assessment of OCD, which until now, have been absent. With the symptom specificity of the OCD-RS and OCD-DOS, clinicians should be able to identify which symptoms have changed over time. The utilization of accurate and sensitive outcome assessment measures could ultimately lead to more efficacious OCD treatment options. In sum, the accuracy-reliability paradigm used in this study contributes to the establishment of assessment tool quality.

*Current Phase*

- Prototypical process for the development of OCD treatment monitoring instruments established by current study

Phase 1

- Single-case design format with children and adolescents who have been diagnosed with OCD
- expand the process utilized in this study to include a OCD symptom checklist similar to the Y-BOCS and a conjoint consultation problem identification interview (PII; Sheridan, Kratochwill, & Bergan, 1996)
- This PII would involve a psychologist or psychiatrist who would facilitate the detailing of the operational definition of the problematic target behaviors as perceived by the parent(s) and/or teacher(s)
- examples and non-example of the target behaviors should be generated

Phase 2

- Use manual developed in this study as a template and train observers to a specified criterion on videotapes to establish accuracy

Phase 3

- Have trained observers use the case specific observation system and rating scale during baseline and treatment phases

  Also, gather traditional assessment scale data (i.e., CY-BOCS, RCMAS)

Figure 8. Developmental phases for future research

Phase 4

- Establish convergent and divergent validity

Phase 5

- Repeated Application of this process to more cases of children and adolescents with OCD.

Phase 6

- Publish Scales and Training Materials

Phase 7

- Extend and validate scales for diagnosis

Phase 8

- Extend process to other observable childhood disorders

Figure 8 (Cont.) Developmental phases for future research

## References

Achenbach, T.M., McConaughy, S.H., Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin, 101, 213-232.

Adams, G.B., Waas, G.A., March, J.S., & Smith, M.C. (1994). Obsessive compulsive disorder in children and adolescents: The role of the school psychologist in identification, assessment, and treatment. School Psychology Quarterly, 9, 274 - 294.

Allen, A.J., Leonard, H.L., & Swedo, S.E. (1995). Case Study: A new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's Syndrome. Journal of American Academy of Child and Adolescent Psychiatry, 34, 307-311.

American Psychiatric Association (1994). Diagnostic and Statistical Manual of Mental Disorders (4th Edition). Washington, DC: Author.

Berg, C.Z. (1989). Behavioral assessment techniques for childhood obsessive-compulsive disorder. In J. L. Rapoport (Ed.) Obsessive-Compulsive Disorder in Children and Adolescents (pp. 41-70). Washington, DC: American Psychiatric Press.

Berg, C.Z., Rapoport, J.L., & Flament, M. (1986). The Leyton Obsessional Inventory-Child Version. Journal of American Academy for Child and Adolescent Psychiatry, 25, 84-91.

Berg, C.Z., Whitaker, A., Davies, M., Flament, M.F., & Rapoport, J.L. (1988). The survey form of the Leyton Obsessional Inventory-Child version. Journal of American Academy of Child and Adolescent Psychiatry, 27, 759 - 763.

Busk, P.L., & Serlin, R.C. (1992). Meta-analysis for single-case research. In T.R. Kratochwill and J.R. Levin (Eds.), Single-case research design and analysis: New directions for psychology and education (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum.

Carter, A.S., Pauls, D. L., & Leckman, J. F. (1995). The development of obsessionality: Continuities and discontinuities. In D. Cicchetti & D.J. Cohen (Eds.), Developmental Psychopathology. Volume 2: Risk, Disorder, and Adaptation (pp. 609-632).

Clarizio, H.F. (1991). Obsessive-compulsive disorder: The secretive syndrome. Psychology in the Schools, 28, 106-115.

Como, P.G. (1997). Neuropsychological tests for obsessive-compulsive disorder and attention deficit hyperactivity disorder. Neurologic Clinics Of North America, 15, 255-265.

Cone, J. D. (1978). The behavioral assessment grid (BAG): A conceptual framework and a taxonomy. Behavioral Therapy, 9, 882-888.

Cone, J.D. (1981). Psychometric considerations. In M. Hersen & A. Bellack (Eds.), Behavioral Assessment: A Practical Handbook (pp. 38-70). New York: Pergamon Press.

Cone, J.D. (1992). Accuracy and curriculum-based measurement. School psychology Quarterly, 7, 22-26.

DeVeaugh-Geiss, J., Moroz, G., Biederman, J., Cantwell, D., Fontain, R., Greist, J.H., Reichler, R., Katz, R., & Landau, P. (1992). Clomipramine hydrochloride in childhood and adolescent obsessive-compulsive disorder-A multicenter trial. Journal of the American Academy of Child and Adolescent Psychiatry, 31, 45-49.

Elliott, S.N., Gresham, F.M., Freeman, T., & McCloskey, G. (1988). Teacher and observer ratings of children's social skills: Validation of the social skills rating scales. Journal of Psychoeducational Assessment, 6, 152-161.

Evans, D.W., Leckman, J.F., Carter, A., Reznick, J.S., Henshaw, D., King, B.A., & Pauls, D. (1997). Rituals, habit, and perfectionism: The prevalence and development of compulsive-like behavior in normal young children. Child Development, 68, 58-68.

Flament, M.F., Rapoport, J.L., Berg, C.A., Sceery, W., Kilts, C., Mellstrom, B., & Linnoila, M. (1985). Clomipramine treatment of childhood obsessive compulsive disorder: A double-blind controlled study. Archives of General Psychiatry, 42, 977-983.

Flament, M.F., Whitaker, A., Rapoport, J.L., Davies, M., Berg, C.Z., Kalikow, K., Sceery, W., & Shaffer, D. (1988). Obsessive compulsive disorder in adolescence: An epidemiological study. Journal of American Academy of Child and Adolescent Psychiatry 27 764 -771.

Foster, S.L., & Cone, J.D. (1995). Validity issues in clinical assessment. Psychological Assessment, 7, 48-260.

Foster, S.L. & Cone, J.D. (1986). In A.R. Ciminero, K.S. Calhoun, & H.E. Adams (Eds.), Handbook of Behavioral Assessment, second edition (pp. 253-325).

Francis, G., & Gragg, R.A. (1997). Childhood obsessive compulsive disorder. Thousand Oaks, CA: Sage publications.

Goodman, W.K. & Price, L.H. (1992). Assessment of severity and change in obsessive compulsive disorder. Psychiatric Clinics of North America, 15, 861-869.

Goodman, W.K., Rasmussen, S.A., Price, L.H., Riddle, M.A., & Rapoport, J.L. (1990). Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) (pp. 1 21). Connecticut Mental Health Center: New Haven, CT.

Hanna, G. (1995). Demographic and clinical features of Obsessive-Compulsive disorder in children and adolescents. Journal of Academic Child and Adolescent Psychiatry, 34, 19 - 27.

Henin, A., & Kendall, P.C. (1997). Obsessive-compulsive disorder in childhood and adolescence. Advances in Clinical Child Psychology, 19, 75-131.

Ilg, F.L., Ames, L.B., & Baker, S.M. (1981). Child Behavior. New York: Harper Collins.

Johnston, H. F. & March, J.S. (1992). Obsessive-Compulsive Disorder in children and adolescents. In W.M. Reynolds (Ed.), Internalizing Disorders in Children and Adolescents (pp. 107-148). New York: John Wiley & Sons.

King, N., Inglis, S., Jenkins, M., Myerson, N. & Ollendick, T. (1995). Test-retest reliability of the survey form of the Leyton Obsessional Inventory-Child version. Perceptual and Motor Skills, 80, 1200-1202.

Kratochwill, T. R., & McGivern, J. E. (1997). Clinical diagnosis, behavioral assessment, and functional analysis: Examining the connection between assessment and intervention. School Psychology Review.

Leonard, H. L. (1995). Update on obsessive-compulsive and tic disorders in children and adolescents. Currents in Affective Illness, 14, 5-13.

Leonard, H., Goldberger, E.L., Rapoport, J.L., Cheslow, B.S., & Swedo, S. E. (1990). Childhood rituals: Normal development or obsessive-compulsive symptoms? Journal of American Academy of Child and Adolescent Psychiatry, 29, 17-23.

Leonard, H.L., Swedo, S.E., March, J.S., & Rapoport, J.L. (1995). Obsessive-compulsive disorder. In C.O. Gabbard (Ed.), Treatment of Psychiatric Disorders, Second Edition (pp. 301-313). Washington, DC: American Psychiatric Press.

Leonard, H.L., Swedo, S.E., Rapoport, J.L., Koby, E.V., Lenane, M.C., Cheslow, D.L., & Hamburger, M.A. (1989). Treatment of OCD with clomipramine and desipramine in children and adolescents: A double-blind crossover comparison. Archives of General psychiatry, 46, 1088-1092.

Marascuilo, L.A., & Serlin, R.C. (1988). Statistical methods for the social and behavioral sciences. New York: W.H. Freeman and Company.

March, J.S., Johnston, H., & Greist, J. (1990). The future of research in obsessive-compulsive disorder. In M. Jenike, L. Baer, & W. Minichello (Eds.), Obsessive-Compulsive Disorder. Littleton, MA: PSG.

March, J.S., & Leonard, H.L. (1996). Obsessive-compulsive disorder in children and adolescents: A review of the past 10 years. Journal of American Academy of Child and Adolescent Psychiatry, 35, 1265-1273.

March, J.S., Leonard, H.L., & Swedo, S.E. (1995). Obsessive-compulsive disorder. In J.S. March (Ed.), *Anxiety Disorders in Children and Adolescents* (pp. 251 - 275). New York: Guilford Press.

Merrell, W. (1993). Using behavior rating scales to assess social skills and antisocial behavior in school settings: Development of the School Social Behavior Scales. *School Psychology Review, 22*, 115-133.

Piacentini, J., Jaffer, M., Gitow, A., Graae, F., Davies, S.O., Del Bene, D., & Liebowitz, M. (1992). Psychopharmacologic treatment of child and adolescent obsessive compulsive disorder. *Psychiatric Clinics of North America, 15*, 87-107.

Querioz, L., Motta, M.A., Madi, M., Sossai, D.L., & Boren, J.J. (1981). A functional analysis of obsessive compulsive problems with related therapeutic procedures. *Behavior Research and Therapy, 19,* 377 -388.

Racine, C.N. (1994). *The relationship between observations and ratings of children's social behavior: An extension of the accuracy-reliability paradigm.* Unpublished manuscript, University of Wisconsin at Madison.

Rapoport, J.L. (1992). Obsessive-Compulsive Disorder. *New Directions for Mental Health services, 54*, 25 - 28.

Rasmussen, S.A. (1994). Genetic studies of obsessive compulsive disorder. In E. Hollander, J. Zohar, D. Marazziti, and B. Olivier (Eds.), Current Insights in Obsessive Compulsive Disorder (pp. 105-114).. New York: John Wiley & Sons.

Reed, M.L. & Edelbrock, C. (1983). Reliability and validity of the Direct Observation Form of the Child Behavior Checklist. Journal of Abnormal Child Psychology, 11, 521-530.

Rettew, D.C., Swedo, S.E., Leonard, H. L., et al. (1992). Obsessions and compulsions across time in 79 children and adolescents with obsessive-compulsive disorder. Journal of the American academy of Child and Adolescent Psychiatry, 31, 1050-1056.

Riddle, M.A., Scahill, L., Smith, J., King, R., Hardin, M., Ort, S., & Epperson, M. (1993). Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS). Biological Psychiatry, 33, A103 -A104.

Robertson, S.L. (1993) The relationship between observations and ratings of children's social behavior. Unpublished master's thesis, University of Wisconsin-Madison, Madison, WI.

Robbins, L.N., Helzer, J.E., Weissman, M.M., Oravashel, H., Gruenberg, E., Burke, J.D., & Regier, D.A.(1984). Lifetime prevalence of specific psychiatric disorders in three sites. Archives of General Psychiatry, 41, 949-959.

Scahill, L., Riddle, M.A., McSwiggin-Hardin, M., Ort, S.I., King, R.A., Goodman, W.K., Cicchetti, D., Leckman, J.F. (1997). Children's yale-brown obsessive compulsive scale: Reliability and validity. Journal of American Academy of Child and Adolescent Psychiatry, 36, 844-852.

Seymour, W.L., & Kratochwill, T.R. (1998). Measuring the function of selective mutism. The selective mutism analysis scale. Manuscript submitted for publication.

Stallings, P. & March, J.S. (1995). Assessment. In. J.S. March (Ed.), Anxiety disorders in children and adolescents (pp. 125-147). New York: Guilford Press.

Swedo, S.E. (1994). Sydenham's chorea: A model for childhood autoimmune neuropsychiatric disorders. Journal of American Medical Association, 272, 1788-1791.

Swedo, S.E., Leonard, H.L., & Rapoport, J.L. (1990). Childhoodonset obsessive-compulsive disorder. In M.A. Jenike, L. Baer, & W.E. Minichiello (Eds.), Obsessive-Compulsive Disorders: Theory and Management, Second Edition (pp. 28-38). Littleton, MA: Year Book Medical Publishers, Inc.

Taylor, S. (1995). Assessment of obsessions and compulsions: Reliability, validity, and sensitivity to treatment effects. Clinical Psychology Review, 15, 261 -296.

VanNoppen, B.L., Pato, M.T., & Rasmussen, S. (1993) Learning to Live with OCD. 3rd Edition. Milford, CT: OC Foundation.

Velhulst, F.C., & Koot, H. M. (1992). Child psychiatric epidemiology: Concepts, methods, and findings. Newbury Park, CA: Sage Publications.

Whitaker, A., Johnson, J., Shaffer, D., Rapoport, J.L., Kalikow, K., Walsh, B.T., Davies, M., Braiman, S., & Dolinsky, A. (1990). Uncommon troubles in young people: Prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. Archives of General Psychiatry, 47, 487495.

Wolff, R.P., & Wolff, L.S. (1991). Assessment and treatment of obsessive-compulsive disorder in children. Behavior Modification, 15, 372-393.

# Appendix A:

# Consent Letters

## INFORMED CONSENT FOR MINOR (Student Actor)

We would like to invite you to be involved in a research study that will help others who may have difficulty in school. Sometimes students much like yourself have a hidden problem where they think about things a lot, and then they feel like they have to do silly stuff to make the bad thoughts go away. For example, a student might think that something bad will happen to their family. In order to make this thought go away, they try things like counting to a certain number. This secret problem is called Obsessive-Compulsive Disorder, or OCD.

We are creating a videotape showing a classroom situation where students act like students and possibly even act like a student who has OCD. We are doing this to try to help teachers and parents of students with OCD because students with OCD sometimes have a hard time in school and doing their work.

You were asked to take part in this study because Mrs. ________ thought you might enjoy acting as a student.

You will have to do two things. First, you and other student actors will need to rehearse with your coach, Caroline Racine. This coaching will take about 2 hours a week for 4 weeks (8 hours total). Second, you will need to agree to be videotaped in a mock classroom situation for about 4 hours on June 9, 1997 at 1 pm. This videotaping will occur at UW-Hospitals and Clinics in Madison.

Even if you start out in the study, you can quit at any time and nothing will happen to you.

We now invite you to sign the bottom of this form indicating that you understand the above information and agree to participate in this study.

----------------------------

I have had this study explained to me in a way I understand, and I have had the chance to ask questions. I understand that I do not have to participate in this study, and that I do not have to give a reason for not participating.

If I sign this form, it means that I agree to participate in this study. My parent(s) will also agree to have me in the study.

Signature of participant________________________________ Date

Signature of investigator________________________________ Date

## Parent Consent for Minor Participation

April 14, 1997

Dear Parent:

We are writing to invite your child to participate in a research project on Obsessive Compulsive Disorder (OCD). The project is designed to create a new rating scale and observation form for educational and professional purposes.

This research project is funded by the U.S. Department of Education, and we are asking for your child's participation (approximately 20 hours). You will be reimbursed for your travel, parking, and any other expenses that you may incur from participating in this project. The research is conducted by Caroline N. Racine, a doctoral student in the Department of Educational Psychology at the University of Wisconsin-Madison and a certified school psychologist; Dr. Thomas R. Kratochwill, Professor of Educational Psychology at the University of Wisconsin-Madison; and Dr. Hugh Johnston, Professor in the Department of Psychiatry at the University of Wisconsin-Madison. This project has received approval from the University of Wisconsin Human Subjects Committee. We believe that there are virtually no risks involved in your child's participation. One foreseeable risk is your child's anxiety about being videotaped. Therefore, the project will be explained to your child in developmentally appropriate language and concerns will be addressed throughout the video process. Finally, the credits at the end of the videotape will state that the students on the videotape were actors.

The specific goal of this project is to develop a high quality method of rating of those children who have OCD. First, we will videotape a classroom situation involving student actors with/without OCD. Your child will need to rehearse with their coach, Caroline Racine for about 2 hours a week for four weeks (i.e., 8 hours total). After the script is rehearsed, your child will need to come to the UW-Hospitals and Clinics in Madison for approximately 4 hours on June 9, 1997. This videotape will be viewed by 12 OCD experts across the country; however, they are aware that the videotape(s) is confidential and must be sent back to UW-Madison. Graduate student observers will also watch the videotape and use the new instruments.

We do not anticipate any drawbacks from your child's participation in this research project. Your child's name will not be made public in any presentations or publications. You will be given the option of keeping a videotape at the completion of the study and deciding if you want your child's name in the credit section of the videotape. You or your child may withdraw from the project at any time without penalty or loss of benefits to you or your child.

If you agree to have your child participate, please sign the attached form. If you have any questions or concerns regarding your participation in the project, please contact Caroline Racine at (608) 238 - 8309 or 262 -3815. Thank you for your time.

Sincerely,

Caroline N. Racine, MS
School Psychologist
Student Investigator
(608) 262-3815
(608) 238-8309

Thomas R. Kratochwill, PhD
Psychologist
Co-Principal Investigator
(608) 262-5912

Hugh Johnston, MD
Psychiatrist
Co-Principal Investigator
(608) 263-6084

# PARENTAL CONSENT FOR THEIR CHILD'S PARTICIPATION IN OCD PROJECT

Caroline N. Racine, MS: Student Investigator
Thomas R. Kratochwill, PhD & Hugh Johnston, MD: Co-Principal Investigators

I, ____________________, acknowledge being informed to my satisfaction of the purpose, benefits, and risks of my child's participation in this project. I also have been informed to my satisfaction of the procedures that will be used in the OCD rating scale development project. It is my understanding that the procedures will involve the following:

1. Allow my child to rehearse with a coach and be videotaped. This experience will involve approximately 20 hours of my child's time over the course of a couple of months.
2. Allow my child's videotape to be viewed by project coordinators, experts in OCD, and graduate student observers.
3. Allow the videotape to be used for training purposes.
4. I will have the opportunity to keep a videotape at the completion of the study.
5. My child and I will have the opportunity to decide if my child's identity shall be placed in the credit section of the videotape.

I understand that confidentiality of my child is assured, and will not be reported in any formal discussion or publication of the project. I understand that I may withdraw my child from the program at any time without penalty or any loss of benefit to me or my child. I further understand that there are virtually no risks associated with my child's participation in the project.

______________________________________________

Child's Name

______________________________________________

Parent Signature Date

*My signature also indicates that I have received a copy of this consent form

## Parental Observation Consent

Dear Parent:

My name is Caroline N. Racine, and I am a doctoral student in Educational Psychology. I am conducting a research project as a part of my dissertation to develop a new rating scale and observation form for children with OCD. In order to develop these new instruments, I need to write a script involving a child who has OCD. In order for this video to be realistic, I need to be able to see how OCD manifests itself in the school setting first hand. Therefore, I am writing you to seek your permission to observe ________________ in his/her classroom. My observation of ______________ will be approximately 2 to 4 hours long, and your child will not be informed that I am observing him/her. Additionally, no one in the class will be informed that I am observing your child. The sole purpose of my observation will be to gather information on how OCD behaviors are displayed in the school setting, if at all. All information collected will remain confidential.

If you have any questions or concerns, please do not hesitate to call me at (608) 238-8309. If you will agree to let me observe your child please sign the bottom of this letter.

Thanks for your time.
Sincerely,

Caroline N. Racine, MS

---

Yes, __________________ can be observed by Caroline N. Racine for
(child's name)
approximately 2 to 4 hours.

________________________________________________________
Parent signature Date

## Observer Consent Letter

October 15, 1996

Name
Address
City, State, Zip

Dear Observer:

We are writing to invite you to participate in a research project that is designed to create a new outcome assessment measure for the purpose of accurately assessing and monitoring change in a child's OCD behaviors during and after treatment. Your participation mayhelp other children who suffer from OCD and may enhance the advancement of assessment and treatment in the field of OCD.

This research project is funded by the U.S. Department of Education, and we are asking for your participation (approximately 6 -10 hours) over the course of 12 months. You will be paid an hourly wage of $8/hr. for your participation in this project. The research is conducted by Caroline N. Racine, a doctoral student in the Department of Educational Psychology at the University of Wisconsin-Madison and a certified school psychologist; Dr. Thomas R. Kratochwill, Professor of Educational Psychology at the University of Wisconsin-Madison; and Dr. Hugh Johnston, Professor in the Department of Psychiatry at the University of Wisconsin-Madison. This project has received approval from the University of Wisconsin Human Subjects Committee. We believe that there are no risks involved in your participation.

The specific goal of this project is to develop a high quality, accurate outcome assessment measure to be used in assessing the treatment of children who have OCD which significantly hinders their school performance. It will be important to create an instrument that can be tailored to a child with OCD on a case-by-case basis since OCD in individuals is idiosyncratic.

Your participation will be needed toward the end of the project and after the rating scale and direct observation record have been created. Again, the time commitment for this project will be approximately 6 - 10 hours. You will be asked to watch a videotape of a child in a school setting, and to use the rating scale and direct observation record as it has been explained to you.

We do not anticipate any difficulties or any risks from participating in this research project. All information collected from your participation will be kept confidential and your identity will not be made public in any presentations or publications . You may withdraw from the project at any time without penalty or loss of benefits to you. Of course, we cannot continue to pay you if you withdraw from the project.

If you agree to participate, please sign the attached form. If you have any questions or concerns regarding your participation in the project, please contact Caroline Racine at 238 - 8309 or 262 -3815. Thank you for your time. We look forward to hearing from you.

Sincerely,

Caroline N. Racine, MS
School Psychologist
Student Investigator
(608) 262-3815
(608) 238 -8309

Thomas R. Kratochwill, PhD
Psychologist
Co-Principal Investigator
(608)262-5912

Hugh Johnston, MD
Psychiatrist
Co-Principal Investigator
(608) 263-6084

## OBSERVER CONSENT FOR PARTICIPATION IN OCD PROJECT

Caroline N. Racine, MS: Student Investigator
Thomas R. Kratochwill, PhD & Hugh Johnston, MD: Co-Principal Investigators

I acknowledge being informed to my satisfaction of the purpose, goals, benefits, risks, and procedures of my participation in the OCD rating scale development project, and I agree to participate. I understand that confidentiality of my identity is assured, and will not be reported in any formal discussion or publication of the project. I agree not to disclose any information about this study to any other person. I understand that I may withdraw from the program at any time without penalty or any loss of benefit to me. I further understand that there are no risks associated with my participation in this project.

Observer Signature Date

## Expert Consent Letter

October. 1997

(Address)

Dear Dr. Expert:

We are writing to invite you to participate in a research project that is designed to create a new outcome assessment measure for the purpose of accurately assessing and monitoring change in a child's OCD behaviors during and after treatment. You have been selected to participate in this project due to your expertise on OCD in children. Your participation may help other children who suffer from OCD and may enhance the advancement of assessment and treatment of OCD.

This research project is funded for a year by the U.S. Department of Education, and we are asking for your participation (approximately 3 hours) over the course of the 12 months. You will be paid an honorarium of $500 for your participation in this project. The research is conducted by Caroline N. Racine, a doctoral student in the Department of Educational Psychology at the University of Wisconsin-Madison and a certified school psychologist; Dr. Thomas R. Kratochwill, Professor of Educational Psychology at the University of Wisconsin-Madison; and Dr. Hugh Johnston, Professor in the Department of Psychiatry at the University of Wisconsin-Madison. This project has received approval from the University of Wisconsin Human Subjects Committee. We believe that there are no risks involved in your participation.

The specific goal of this project is to develop a high quality, accurate outcome assessment measure to be used in assessing the treatment of children who have OCD which significantly hinders their school performance. It will be important to create an instrument that can be tailored to a child with OCD on a case-by-case basis since OCD in individuals is idiosyncratic.

Currently, your participation is needed because the rating scale and direct observation system have been created. The time commitment for this project will be approximately 3 hours. You are asked to watch a videotape of a child in a school setting and to use the rating scale and direct observation system as it has been explained to you in the manual provided. You are required to send back the signed consent form, videotape, the rating scale, and the direct observation record in the return, postage paid envelope that is enclosed for your convenience. If you choose not to participate, you still must send back all the materials . You will be **paid $500** as an incentive to participate and return all materials promptly. You must send all the materials along with the completed tasks back by **December 15, 1997** to ensure proper payment. We request confidentiality of the videotape.

We do not anticipate any difficulties or any risks from participating in this research project. All data collected from your participation will be kept confidential, and your identity will not be made public in any presentations or publications. You may withdraw from the project at any time without penalty or loss of benefits to you.

If you agree to participate, please sign the attached form and complete the direct observation system and rating scales as instructed. If you have any questions or concerns regarding your participation in the project, please contact Caroline Racine at (608) 238 - 8309 or (608) 262 -3815 or via email (cnracine@students.wisc.edu). Thank you for your time. We look forward to receiving your materials.

Sincerely,

Caroline N. Racine, MS
School Psychologist
Student Investigator
(608) 238-8309
(608) 262-3815

Thomas R. Kratochwill, PhD
Psychologist
Co-Principal Investigator
(608) 262-5912

Hugh Johnston, MD
Psychiatrist
Co-Principal Investigator
(608) 263-6084

*Enclosures*
- *expert consent form*
- *videotape*
- *manual*
- *OCD-DOS & RS*
- *return checklist*
- *return envelope*

## EXPERT CONSENT FOR PARTICIPATION IN OCD PROJECT

Caroline N. Racine, MS: Student Investigator
Thomas R. Kratochwill, PhD. & Hugh Johnston, MD: Co-Principal Investigators

I acknowledge being informed to my satisfaction of the purpose, goals, benefits, risks, and procedures of my participation in the OCD rating scale development project: and I agree to participate. I understand that confidentiality of my identity is assured, and will not be reported in any formal discussion or publication of the project. I understand that I may withdraw from the program at any time without penalty or any loss of benefit to me. Of course, we cannot pay you if you choose not to participate. I further understand that there are no risks associated with my participation in this project. I will send back the videotape and instruments in a prompt manner.

Expert Signature Date

Address to send honorarium:

Social Security #:

## Teacher Consent Form (Actress)

April 14, 1997

Dear Mrs. __________:

We are writing to invite you to participate in a research project that is designed to create a new outcome assessment measure for the purpose of accurately assessing and monitoring change in a child's OCD behaviors during and after treatment. Your participation may help other children who suffer from OCD and may enhance the advancement of assessment and treatment in the field of OCD.

This research project is funded by the U.S. Department of Education, and we are asking for your participation (approximately 20 hours) over the course of 12 months. You will need to practice approximately 2 hours per week for four weeks. After the script is rehearsed, you will need to come to UW Hospitals and Clinics in Madison for approximately four hours on June 9, 1997 at 1pm. Specifically, we would like you to act as the role of teacher in a videotape we are creating. You will be paid a stipend of $600 for your participation in this project. The research is conducted by Caroline N. Racine, a doctoral student in the Department of Educational Psychology at the University of Wisconsin-Madison and a certified school psychologist; Dr. Thomas R. Kratochwill, Professor of Educational Psychology at the University of Wisconsin-Madison; and Dr. Hugh Johnston, Professor in the Department of Psychiatry at the University of Wisconsin-Madison. This project has received approval from the University of Wisconsin Human Subjects Committee. We believe that there are no risks involved in your participation.

The specific goal of this project is to develop a high quality, accurate outcome assessment measure to be used in assessing the treatment of children who have OCD which significantly hinders their school performance. It will be important to create an instrument that can be tailored to a child with OCD on a case-by-case basis since OCD in individuals is idiosyncratic.

We do not anticipate any difficulties or any risks from participating in this research project. All information collected from your participation will be kept confidential and your identity will not be made public in any presentations or publications (e.g., your name or identity will not be used). You will be given the opportunity to decide if you would like your name to appear in the credit section at the end of the videotape when the videotape has been completed. You may withdraw from the project at any time without penalty or loss of benefits to you. Of course, we cannot continue to pay you if you withdraw from the project.

If you agree to participate, please sign the attached form. If you have any questions or concerns regarding your participation in the project, please contact Caroline Racine at 238 - 8309 or 262 -3815. Thank you for your time. We look forward to hearing from you.

Sincerely,

Caroline N. Racine, MS
School Psychologist
Student Investigator
(608) 262-3815
(608) 238 -8309

Thomas R. Kratochwill, PhD
Psychologist
Co-Principal Investigator
(608)262-5912

Hugh Johnston, MD
Psychiatrist
Co-Principal Investigator
(608) 263-6084

## ADULT ACTOR CONSENT FOR PARTICIPATION IN OCD PROJECT

Caroline N. Racine, MS: Student Investigator
Thomas R. Kratochwill, PhD & Hugh Johnston, MD: Co-Principal Investigators

I acknowledge being informed to my satisfaction of the purpose, goals, benefits, risks, and procedures of my participation in the OCD rating scale development project, and I agree to participate. I understand that I may withdraw from the program at any time without penalty or any loss of benefit to me. I further understand that there are no risks associated with my participation in this project.

Participant Signature　　　　　　　　　　　　　　　　Date

# Appendix B:

## Behavior Codes

## OCD-RS

## OCD-DOS

# *Obsessive-Compulsive Disorder-Rating Scale*
# *for children and adolescents*
## *OCD-RS*

### Student Information

| | |
|---|---|
| Name: | Date of Birth: |
| School: | Grade: |
| Referred by: | Age: |

### Rating Information

| | |
|---|---|
| Date of Rating: | Rater: |
| Class: | Teacher: |

Type of Setting (i.e., large group, small group, independent seatwork, etc.):

### Additional Information/Notes

## OCD-RS Behavior Codes

### Reassurance Seeking (RS)

This type of behavior includes excessive questioning of directions, assignments, activities, time limits, classroom rules, etc. Basically, this person's behavior conveys a need for reassurance by someone else. RS is distinguished from typical questions primarily by quantity. These RS questions are different from that of clarification questions in that the listener/observer has the intuitive sense of what the answer is.

### Writing Rituals (W)

This behavior consists of writing, rewriting and/or retracing. Difficulty with writing could include the following behaviors: elaborate correction, writing/rewriting, tracing certain letters, and starting whole assignment over.

### Repeating (R)

This behavior is defined as having to do particular actions over and over again. Repeating behaviors include but are not limited to the following: getting school work/supplies out of and putting them back into backpack, having to come through the door several times in a certain way, blowing on fingers before one picks up a pencil or turns a page, standing up and sitting down, mentally counting to a certain number or tapping a certain number of times before starting work, touching a glass to one's lips a certain amount of times, clicking the pen a number of times, reading over and over One episode of repeating is coded as one event. That is, if the student engages in R by standing up and sitting down three times, the observer would code this behavior as one (1) event.

### Symmetry (SY)

This behavior conveys a need for exactness or order. Symmetry behaviors could include the following: lining up a book or paper with the edge of a desk, making sure a writing implement is placed exactly in the center of the desk, arranging and rearranging supplies. One episode of repeating is coded as one event. That is, if the student engages in SY by lining up her book with the edge of the desk several times before stopping, the observer would code this behavior as one (1) event.

# OCD RATING SCALE (OCD-RS)

## *Child and Adolescent Form*

Background Information

This rating scale is designed to measure how often a student exhibits specific, observable OCD behaviors in your classroom and how these specific behaviors interfere with success in your classroom. Specific OCD behaviors will most likely vary across students; therefore, you will be rating those behaviors that have been agreed upon by you, the student's parent(s), and the school psychologist and/or psychiatrist. In addition to rating these behaviors, space is provided for you to describe any additional OCD behaviors that seem to be problematic for the student, should they arise. Your additional observations will be important for intervention monitoring given that OCD symptoms tend to wax and wane over time.

Directions

For each behavior, decide how often the student performed the behavior described.

If student **never** did this behavior, circle 0.

If student **sometimes** did this behavior (i.e., approx. 1-10 times/rating period),-circle 1.

If student **very often** did this behavior (i.e., approx. 11+ times/rating period), circle 2.

For each behavior, you also should rate how much each of these behaviors seemed to interfere with the student's classroom/home/community functioning.

If the behavior did not seem to interfere with classroom/home/community performance and/or did not seem to cause the student distress when he/she is not allowed to perform the behavior, circle 0 (**No Interference**).

If the behavior somewhat interfered with school/home/community performance and/or looked as though the student is in some distress when asked to stop the behavior, circle 1 (**Sometimes Interferes**).

If the behavior interferes with school/home/community performance and/or student was in obvious distress if asked to stop behavior, circle 2 (**Interferes Most of the Time**)

# OCD Rating Scale (OCD-RS)

| OCD Observable Behaviors | How Often? | | | How Much Interference? | | |
|---|---|---|---|---|---|---|
| | Never | Sometimes | Very Often | None | Sometimes | Most of time |
| 1. Reassurance Seeking | 0 | 1 | 2 | 0 | 1 | 2 |
| 2. Writing Rituals | 0 | 1 | 2 | 0 | 1 | 2 |
| 3. Repeating | 0 | 1 | 2 | 0 | 1 | 2 |
| 4. Symmetry | 0 | 1 | 2 | 0 | 1 | 2 |
| 5. | 0 | 1 | 2 | 0 | 1 | 2 |
| 6. | 0 | 1 | 2 | 0 | 1 | 2 |
| 7 | 0 | 1 | 2 | 0 | 1 | 2 |
| 8. | 0 | 1 | 2 | 0 | 1 | 2 |
| 9. | 0 | 1 | 2 | 0 | 1 | 2 |
| 10. Other: | 0 | 1 | 2 | 0 | 1 | 2 |

Additional Notes/Observations:

# Obsessive-Compulsive Disorder- Direct Observation System
# for children and adolescents
# OCD-DOS

## Student Information

Name: Date of Birth:
School: Grade:
Referred by: Age:

## Observation Information

Date of Observation: Observer:
Class: Teacher:
Time Started: Time Ended:
Type of Setting (i.e., large group, small group, independent seatwork, etc.):

## Additional Information/Notes

# OCD - Direct Observation System (OCD-DOS)

## *Adolescent and Child Form*

## Record Form

Directions: This observation system records the frequency of the specified OCD target behavior. Record a tally mark for each occurrence of the OCD behavior under the appropriate columm. Each column corresponds to a five minute interval. For instance, if you observe Reassurance Seeking at 10 minutes and 27 seconds you would record a tally mark in the 3rd column marked 15 because this includes the behavior that occurs during 10:01 and 15:00. For OCD video observation, use intervals as indicated.

| Minute | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | Total |
|---|---|---|---|---|---|---|---|---|---|
| **OCD video #1** | **20:00-25:00** | **25:01-30:00** | **30:01-35:00** | **35:01-40:00** | **40:01-45:00** | | | | |
| 1 Reassurance Seeking | | | | | | | | | |
| 2 Writing Rituals | | | | | | | | | |
| 3 Repeating | | | | | | | | | |
| 4 Symmetry | | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |
| 10 Other | | | | | | | | | |

Additional Observations (i.e., new compulsions observed, antecedents, consequents, sequentioal conditions, etc. ):

## OCD - Direct Observation System (OCD-DOS)

### *Adolescent and Child Form*

### **Record Form**

Directions: This observation system records the frequency of the specified OCD target behavior. Record a tally mark for each occurrence of the OCD behavior under the appropriate columm. Each column corresponds to a five minute interval. For instance, if you observe Reassurance Seeking at 10 minutes and 27 seconds you would record a tally mark in the 3rd column marked 15 because this includes the behavior that occurs during 10:01 and 15:00. For OCD video observation, use intervals as indicated.

| Minute | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | Total |
|---|---|---|---|---|---|---|---|---|---|
| **OCD video #2** | **1:23-6:23** | **6:24-11:23** | **11:24-16:23** | **16:24-21:23** | **21:24-26:23** | | | | |
| 1 Reassurance Seeking | | | | | | | | | |
| 2 Writing Rituals | | | | | | | | | |
| 3 Repeating | | | | | | | | | |
| 4 Symmetry | | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |
| 10 Other | | | | | | | | | |

Additional Observations (i.e., new compulsions observed, antecedents, consequents, sequentioal conditions, etc. ):

## OCD-DOS Behavior Codes

**Reassurance Seeking (RS)**
This type of behavior includes excessive questioning of directions, assignments, activities, time limits, classroom rules, etc. Basically, this person's behavior conveys a need for reassurance by someone else. RS is distinguished from typical questions primarily by quantity. These RS questions are different from that of clarification questions in that the listener/observer has the intuitive sense of what the answer is.

**Writing Rituals (W)**
This behavior consists of writing, rewriting and/or retracing. Difficulty with writing could include the following behaviors: elaborate correction, writing/rewriting, tracing certain letters, and starting whole assignment over.

**Repeating (R)**
This behavior is defined as having to do particular actions over and over again. Repeating behaviors include but are not limited to the following: getting school work/supplies out of and putting them back into backpack; having to come through the door several times in a certain way, blowing on fingers before one picks up a pencil or turns a page, standing up and sitting down, mentally counting to a certain number or tapping a certain number of times before starting work, touching a glass to one's lips a certain amount of times, clicking the pen a number of times, reading over and over.

**Symmetry (SY)**
This behavior conveys a need for exactness or order. Symmetry behaviors could include the following: lining up a book or paper with the edge of a desk, making sure a writing implement is placed exactly in the center of the desk, arranging and rearranging supplies.

# Appendix C

## Training Manual

# Rating Scale and Direct Observation System for Obsessive-Compulsive Disorder in Children and Adolescents

## RATING SCALE AND OBSERVATION MANUAL

Caroline N. Racine, MS
Thomas R. Kratochwill, PhD
Hugh Johnston, MD
University of Wisconsin-Madison

1997

## Section One: Review of Obsessive-Compulsive Disorder

### Introduction

Obsessive-Compulsive Disorder (OCD), originally thought to be a rare disorder, afflicts more than 2% of the population (Whitaker et al., 1990). Among adults who suffer from OCD, approximately one-third to one-half developed OCD when they were children or adolescents (March, Leonard, & Swedo, 1995). Realistically, OCD probably affects more individuals than current epidemiological studies report due to the secrecy of this disorder. Due to the hidden nature of OCD, it is reasonable to assume that it is an under-reported and under-identified disorder that, for a large majority of people, has origins in childhood. In many cases, early detection of this disorder has played a vital role in children's recovery from OCD (Querioz, Motta, Madi, Sossai, & Boren, 1981). Therefore, knowledge and identification of OCD by school psychologists is important to the alleviation of this disorder (Adams, Waas, March, & Smith, 1994).

Briefly, obsessions can be described as "persistent ideas, thoughts, impulses, or images that are experienced as intrusive and inappropriate and that cause marked anxiety or distress" (American Psychiatric Association, 1994, p. 418). Some typical childhood obsessions are fear of contamination, fear of something happening to themselves or loved ones (i.e., fire, death, illness), and need for exactness (Swedo, Leonard, & Rapoport, 1990). Compulsions can be defined as "repetitive behaviors (e.g., hand washing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) the goal of which is to prevent or reduce anxiety or distress, not to provide pleasure or gratification" (American Psychiatric Association, 1994, p. 418). The obsessions and/or compulsions of children with OCD differ from developmentally appropriate obsessive-compulsive behaviors in respect to timing, content, and severity. That is, the obsessions and compulsions of children with OCD usually appear after the developmentally appropriate time, seem odd, and produce an inability to function (March, Leonard, & Swedo, 1995).

## Overview of OCD

According to the Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV) (1994) an individual must have five diagnostic characteristics to meet the criteria for clinical OCD. These characteristics, as outlined by the DSM-IV (1994), are as follows:

A. Either obsessions or compulsions:

*Obsessions* as defined by (1), (2), (3), and (4):
(1) recurrent and persistent thoughts, impulses, or images that are experienced, at some time during the disturbance, as intrusive and inappropriate and that cause marked anxiety or distress
(2) the thoughts, impulses, or images are not simply excessive worries about real-life problems
(3) the person attempts to ignore or suppress such thoughts, impulses, or images, or to neutralize them with some other thought or action
(4) the person recognizes that the obsessional thoughts, impulses, or images are a product of his or her own mind (not imposed from without as in thought insertion)

*Compulsions* as defined by (1) and (2):
(1) repetitive behaviors (e.g., hand washing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) that the person feels driven to perform in response to an obsession, or according to rules that must be applied rigidly
(2) the behaviors or mental acts either are aimed at preventing or reducing distress or preventing some dreaded event or situation; however, these behaviors or mental acts either are not connected in a realistic way with what they are designed to neutralize or prevent or are clearly excessive

B. At some point during the course of the disorder, the person has recognized that the obsessions or compulsions are excessive or unreasonable. **Note**: This does not apply to children.

C. The obsessions or compulsions cause marked distress, are time consuming (take more than 1 hour a day), or significantly interfere with the person's normal routine, occupational (or academic) functioning, or usual social activities or relationships.

D. If another Axis I disorder is present, the content of the obsessions or compulsions is not restricted to it.

E. The disturbance is not due to the direct physiological effects of a substance (e.g., a drug of abuse, medication) or a general medical condition. (p. 422-423).

OCD is diagnosed along Axis I and is considered to be an anxiety disorder Childhood-onset OCD must meet the same criteria as specified for adults in the DSM-IV; however, children may or may not have insight into their obsessions and/or compulsions.

Clinical Course of OCD

The topography of childhood-onset OCD varies on an individual basis as well as intra-individually across time. In other words, the course of childhood-onset OCD does not seem to follow any typical developmental trajectory nor are certain obsessions and/or compulsions predictive of a certain developmental pathway. In a prospective follow-up study by Rettew, Swedo, Leonard, Lenane and Rapoport (1992), qualitative information regarding OCD symptoms across time was reported for a sample of children and adolescents (N=79). Many interesting findings were reported: (a) no significant age-related trends were discovered in terms of symptom type; (b), washing and checking compulsions were exhibited by 47% of the sample at some point during the course of their disorder; (c) symptom constellation changed over time and varied on content and severity dimensions; and (d) children 6-years old and younger were more likely to have compulsions. These authors also noted that pure obsessive OCD was observed less frequently than pure compulsive OCD. The age of onset data from this sample yielded an early onset type and an adolescent onset type. Early onset children were twice as likely as adolescent onset type to be male and to have had a positive family history of OCD or Tourette's Syndrome.

## Prevalence of OCD

Prevalence estimates of OCD usually vary between 2% and 3% of the general population. For example, the Epidemiology Catchment Area (ECA) study reported a 2.5% of lifetime prevalence in the population at large (Robins et al., 1984). Flament and colleagues (1988) conducted a survey of 5,596 high school students, and 20 adolescents were identified as having OCD (0.4%). When the sample was weighted, the estimated prevalence rate for adolescents with OCD was 1%. Additionally, the majority of childhood and adolescent OCD studies indicate that most children show OCD symptoms during 10-14 years of age and another one-third report symptoms before 9 years of age (Riddle et al., 1990; Swedo et al., 1989). The true prevalence of childhood-onset OCD is difficult to ascertain given that there have been no appropriate epidemiological child and adolescent studies (Carter, Pauls, & Leckman, 1995).

## Complications from OCD

Complications from OCD can affect children and adolescents both at school and home.

Impact of OCD at school. Although the prevalence rate of OCD in children and adolescents is not extraordinary, it is believed to be 20 to 40 times more common than previously reported (Clarizio, 1991). The secretive nature of this problem makes it difficult to identify those who have OCD. Although there has been no systematic investigation as to how youth with OCD function in school. it is important for school personnel to be knowledgeable about OCD and be able to critically evaluate referrals for special education. Due to the individual variability of OCD across students. OCD will impact students in different situations. Among the most common school problems that students with OCD experience are: (a) poor peer relations; (b) high absentee rates; (c) a decline in scholastic aptitude and overall school functioning; and (d) an increase in drug and alcohol abuse (Adams et al., 1994; Clarizio, 1991). Many of these students also are referred and placed in special education due to their inability to function in a regular education classroom; however, other students with OCD remain in the regular education setting.

Impact of OCD at home. Each family who has a child or adolescent with OCD copes differently. VanNoppen, Pato, and Rasmussen (1993) delineate five different coping strategies typically adopted by families who have a family member with OCD. First, there are those families who help with the rituals in an effort to maintain the status quo. Second, there are those families who do not assist in the rituals, but allow the compulsions to occur. Third, there are those families who ignore or fail to acknowledge the obsessions and compulsions. Fourth, many families will respond differentially to the family member with OCD. Finally, there are those families who demonstrate inconsistent responses to the family member with OCD. Children and adolescents with OCD often try hard to keep their OCD behaviors hidden from other family members. Often times, children and adolescents feel isolated and ashamed because they realize that their obsessions and/or compulsions are irrational. Children and adolescents with OCD typically are dependent on their parents for many of their needs. This dependence often is troublesome for adolescents because OCD hinders the separation and individuation process that most adolescents experience during the teenage years. The key to coping with a child or adolescent who has OCD is to seek help, allow for open family discussion about OCD, and provide consistent discipline and routines (VanNoppen et al., 1993).

## Associated (Comorbid) Disorders

The comorbidity of OCD with other disorders is very common. In fact, child and adolescent studies that report on comorbidity of their samples reveal similar data in regard to OCD and associated disorders (i.e., Swedo et al., 1989; Swedo, Leonard, & Rapoport, 1992; and Riddle et al., 1990). Data from Swedo and colleagues (1989, 1992) reveal that only 25% of the sample (N = 140) had OCD as their sole diagnosis. The data (N=70) from Swedo et al. (1989) provide an excellent illustration of psychiatric disorders commonly associated with OCD: tic disorder (30%), major depression (26%), specific developmental disorder (24%), simple phobia (17%), overanxious disorder (16%), adjustment disorder with depressed mood (13%), oppositional defiant disorder (11 %), attention deficit disorder (10%), conduct disorder (7%), and separation anxiety disorder (7%). These results highlight the need to consider comorbid conditions when diagnosing children and adolescents with OCD as dual diagnoses will affect intervention planning.

Etiology of OCD

Currently, there are many different hypotheses as to the cause of OCD. The only definitive statement that can be offered with any confidence is that the cause of OCD remains nebulous. Nonetheless, there are various models and hypotheses of OCD which help clinicians to pursue various treatment options. Several contemporary models of OCD are presented briefly to illustrate the scope and diversity of current explanations of OCD. Most of these models have focused primarily on adults; consequently, researchers are left to infer that these adult models also apply to children and adolescents

Neurobiological models. Although there are several neurobiological and neuropsychological models of OCD currently being investigated, the serotonin hypothesis has received the most attention. Evidence that serotonin, a neurotransmitter, is involved in OCD was first obtained when researchers discovered that Clomipramine (CMI), a tricyclic antidepressant as well as a serotonin (and other monoamine) reuptake inhibitor (SRI), reduced OCD symptoms in adults. This finding has been replicated with children and adolescents ( i.e., Flament et al., 1985; Leonard et al., 1989; and DeVeaugh-Geiss et al., 1992). Additional evidence that implicates serotonin dysfunction in the etiology of OCD has been obtained in: (a) direct assays of serotonin and its metabolites; (b) drug challenge studies; and (c) the efficacy of selective serotonin reuptake inhibitors (SSRIs) (Johnston & March, 1992). Research also has demonstrated that individuals with OCD have reduced levels of serotonin in the synaptic cleft as well as post-synaptic hyperactivity. Thus, medications that increase serotonin in the synapse by blocking its reuptake into the presynaptic neuron have alleviated some OCD symptoms (Johnston & March, 1992; Piacentini et al., 1992). This model continues to remain viable; however, it does not account for the sole cause of OCD simply because some individuals do not improve when taking CMI or SSRIs.

Recently, researchers have been investigating the possibility that some forms of childhood-onset OCD may be caused by a viral infection. In reviewing the literature and clinical cases, these investigators discovered that two general patterns of OCD emerged among children and adolescents (Allen et al., 1995; Leonard, 1995; Swedo, 1994). Some children had an insidious onset of OCD that waxed and waned over time. Other children experienced a sudden onset of OCD that remitted between OCD episodes. For this latter group of children, there appeared to be an association between OCD onset and Group A Beta Hemolytic Streptococcal infection, or "strep throat" (Allen et al., 1995; Leonard, 1995; Swedo, 1994). These researchers have likened this process to that of Sydenham's chorea, a form of rheumatic fever, in that, ". . . some patients are genetically predisposed to marshal an autoimmune response to Group A Beta Hemolytic Streptoccus; when they form autoantibodies that are misdirected at various parts of their bodies, such as the heart (which can result in rheumatic carditis) or the brain (which can give rise to chorea)" (Leonard, 1995, p. 13). Hence, Sydenham's chorea might serve as a medical model for OCD. Although this model is promising, more research is needed.

Given the genetic transmission of possible serotonin dysfunction and autoimmune responses, it is not surprising that genetic factors are hypothesized to play a role in the development of OCD as well. In addition to this evidence, family and twin studies have been conducted to provide evidence of OCD heritability. Rettew et al. (1989) (N = 70) reported that 25% of those children and adolescents with OCD had a positive family history of OCD. Most of the family and twin studies report that OCD is familial; however, these studies are not without their methodological weaknesses (Rasmussen, 1994).

Behavioral model. Some believe OCD is a learned behavior that has resulted from classical conditioning and operant learning. This theory is referred to as a two-factor conditioning model (Johnston & March, 1992). For example, a new obsession (i.e., unconditioned stimulus) occurs, the individual experiences anxiety, and then the individual performs a behavior which the individual thinks will reduce this anxiety (i.e., compulsion). This unconditioned stimulus becomes a conditioned stimulus, and operant learning results. Environmental events and cues that surround the conditioned behavior often prompt other behaviors to arise, a process commonly referred to as stimulus generalization. Given that behavior therapy has been successful in symptom alleviation for some individuals, this behavioral model also remains a reasonable explanation as to the cause of OCD.

## Section Two: Assessment of OCD

### Introduction

A best practice approach to the clinical diagnosis of OCD involves the collection of information using a variety of methods and informants across settings (Achenbach, McConaughy, & Howell, 1987). In the assessment of any anxiety disorder, instruments should: (a) be reliable and valid across multiple symptom domains; (b) differentiate symptom clusters; (c) assess the severity of the problem; (d) make use of multiple informants; and (e) be sensitive to behavior change (Starlings & March, 1995). Stallings and March (1995) also assert that assessment devices should facilitate communication among professionals especially in light of the current trend toward a multi-disciplinary team approach to assessment. Assessment purpose also should be delineated and considered prior to the actual assessment. For example, a clinician should reflect on whether the purpose of assessment is for diagnosis, treatment planning, and/or treatment monitoring (Foster & Cone, 1995; Kratochwill & McGivern, 1997). The delineation of assessment purpose will help guide the clinician in determining what type of assessment instrument to use (see Figure 1 for an example). Although, OCD has a controversy-free nosology, OCD presents itself with considerable inter-individual heterogeneity (March, Johnston, & Greist, 1990; Rettew et al., 1992). Thus, practitioners using various assessments must make a concerted effort to account for symptom

|  | RATING SCALE | DIRECT OBSERVATION |
|---|---|---|
| **SYMPTOM CONSTELLATION**<br>**Purpose**<br>**1. Treatment outcome monitoring**<br>**2. Diagnosis** | A rating scale format where OCD observable behavior can be ranked on certain dimensions<br>Example<br>See OCD-Rating Scale (Appendix A) | Event and/or duration recording completed by trained observers. Operational Definitions should be Parallel to Rating Scale Behaviors.<br>Example<br>OCD Behavior — Occurrence — Duration<br>1. tracing & retracing letters — 1 — 10 min<br>2 — 5 min<br>8 min |
| **FUNCTION**<br>**Purpose**<br>**1. Treatment Design** | Rate or Estimate antecedents, consequents, and sequential events when OCD behaviors occur.<br>Example<br>*Child Report*<br>1. How often do you trace letters or words over and over because of avoiding your work?<br>*Teacher Report*<br>1. Does the student seem to trace and retrace words or letters (or numbers) in order to avoid work completion? | Observe antecedents, consequents, and sequential events when OCD behaviors occur.<br>Example<br>A — B — C<br>target behavior ex<br>tracing letters over and over — group lesson — tracing tapping foot — avoids turning in work or answering Qs |

Figure 1. Conceptual Matrix for the Development of Assessment Instruments by Purpose

variability within OCD (Starlings & March. 1995). Clinicians also must consider comorbid conditions as this will impact choice of treatment.

The methods most commonly used by clinicians to assess OCD include the following: a diagnostic interview (structured. semi-structured, and/or unstructured), rating scales (self and/or other), self-monitoring (e.g., behavioral diary method), and direct behavioral observations (Taylor. 1995; Wolff & Wolff, 1991). Due to time efficiency and cost effectiveness, the most relied upon method in clinical practice is most likely the completion of rating scales. Given the focus of this assessment manual, rating scale methods and direct observations will be reviewed.

## Rating Scale Assessment

Given the widespread use of rating scales among clinicians, it is important to be cognizant of the technical adequacies of the most commonly used rating scales. The two most commonly used rating scales, the Leyton Obsessional Inventory-Children's Version (LOI-CV) and the Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) are reviewed in depth. Unfortunately, content validity and accuracy of these scales have not been documented. Consequently, traditional estimates of validity are reported. Both of these scales are purported to have multiple roles in the assessment of OCD; however, they are most commonly used for diagnosis, treatment design, and treatment monitering.

*The Leyton Obsessional Inventory-Child Version (LOI-CV).* The LOI-CV (Berg, Rapoport, & Flament, 1986) is a downward extension of the adult LOI (Cooper & McNeil, 1968) which was designed to assess obsessionality in "houseproud" homemakers. Cooper (1970) adapted it for the assessment of OCD. The LOI-CV, like the LOI, is a 44-item, self-report measure that allows assessment of the following symptom domains: thoughts, checking, dirt and contamination, dangerous objects, cleanliness and tidiness, school work, order and routine, overconscientious, indecision, hoarding, meanness, and magic games. The LOI-CV involves a card-sorting task where the child is presented a symptom on a card, and then he/she drops the symptom card in a "yes" or a "no" box. Positive responses are re-administered to assess the dimensions of resistance (i.e., how hard it is to stop - 5-point Likert-type response format) and interference (i.e., how much symptom interferes with daily functioning- 4- point Likert-type response format). Total symptom scores are calculated by adding the number of positive symptoms cards; and resistance and interference scores are added according to their point values. There also is a 20-item LOI-CV using the same format that has been developed for epidemiological studies.

In terms of reported technical adequacies for the LOI-CV (44 item), data have been scant. In fact, a literature search conducted by the Obsessive-Compulsive Information Center at the Dean Foundation located in Madison, Wisconsin revealed only 19 articles that referenced the LOI-CV. Of these 19 articles, only one addressed the technical adequacies of this instrument. Berg, Rapoport, and Flament (1986) specifically tested the reliability and validity of the LOI-CV by studying three groups: (1 ) 26 adolescents (age range = 10-18 yrs.; M= 14.3 yrs.; 17 males, 9 females) who were diagnosed with OCD; (2) 28 adolescent controls (age range = 11-18 yrs.; M = 13.7 yrs.; 22 males, 6 females), who were matched for age, and (3) 14 psychiatric adolescents (age range = 10-17 yrs.; M = 13.2; 10 males, 4 females) who had some obsessive symptoms, but OCD was not their primary diagnosis. Of the first group, 19 participated in a double-blind crossover drug treatment trial. A one-way ANOVA found significant group differences resulted for the number of "yes," resistance, and interference scores. The test-retest reliability of 10 patients across a placebo phase of 5 weeks demonstrated intraclass coefficients of .96, .97, and .94 for the "yes," resistance, and interference scores, respectively. Additionally, Pearson product-moment correlations were computed between the LOI-CV and other OCD rating scales. Correlations ranged from .77 to .89 for both the LOI-CV and the Obsessive Compulsive Rating (OCR, Rapoport et al., 1980), the Comprehensive Psychopathological Rating Scale-Obsessive Compulsive Subscale (CPRS-OC; Thoren et al., 1980), and the NIMH Obsessive Compulsive scale (NIMH-OC; Rapoport et al., unpublished). Correlations between the LOI-CV and the NIMH Global subscale (NIMH-G,Murphy et al., 1982) ranged from .69 to .77. Paired t-tests were also computed between the end of baseline and the end of treatment scores. Significant differences were reported, indicating that the LOI-CV can provide useful information about clinical change.

Three studies from the same literature search were identified for providing information on the LOI-CV (20 items) epidemiological version. Flament and colleagues (1988) gathered epidemiological information from 356 high school students. These researchers reported the sensitivity of the LOI-CV (20 items) as 75%, the specificity as 84%, and the predictive validity as 18%. The sensitivity percentage indicates the number of cases that test positive for the target disorder, and the specificity percentage means the number of cases that test negative for the target disorder. Verhulst and Koot (1992) assert that these measures of validity do not demonstrate the intrinsic properties of an instrument. In fact, these percentages will vary between samples. The predictive value indicates the likelihood that a positive test result will predict a certain disorder (Verhulst & Koot, 1992). Given the 18% predictive validity, Flament and researchers (1988) assert that the LOI-CV (20 items) is an adequate screening measure. Berg and colleagues (1988) reported an internal consistency estimate of .81 based on Cronbach's $\alpha$ coefficient statistic. King, Myerson, Inglis, Jenkins, and Ollendick (1995) examined the reliability of the LOI-CV (20 items) in an Australian sample (N=1602) with an age range of 8 to 16 years. These researchers reported a Cronbach's $\alpha$ of .76. Two-week test-retest reliability was calculated using Total Obsessive scores and Total Interference scores for three different age cohorts (8-10 yr.-olds, 11-13 yr.-olds, and 14-16 yr.- olds) on 106 randomly selected participants. For the youngest cohort, .51 and .65 on Total Obsessive and Total Interference scores were calculated, respectively. For the middle cohort, .75 and .81, for Total Obsessive and Total Interference scores were derived, respectively. For the oldest cohort, .83 and .57 for Total Obsessive and Total Interference scores were found, respectively. Total Obsessive scores were observed to increase with age; however, this was not true for the Interference scores. (see Table 1 for Technical Information on LOI-CV.)

This dearth of tested reliability and validity for the LOI-CV (44-item and 20-item scales) makes it extremely hard to judge its technical adequacy as well as usefulness in the assessment process. Additionally, the studies that were found are difficult to compare given the lack of common statistics used to judge the instrument. Finally, it should be underscored that traditional validity studies are distinctly different from epidemiological validity studies, and caution should be rendered when trying to compare statistics across studies. Given these caveats about the lack of reported technical adequacies, the advantages to the LOI-CV are: (a) it is the only self-report measure available for children; (b) it provides a measure of presenting symptoms, resistance, and interference; and

Table 1

Research studies investigating psychometric properties of the LOI-CV and CY-BOCS

| RESEARCH STUDY | RELIABILITY | | | VALIDITY | | |
|---|---|---|---|---|---|---|
| | Internal Consistency | Test-Retest | Inter-Rater | Content | Criterion | Construct |
| LOI-CV (44) | | | | | | |
| Berg et al , 1986<br>(N = 54) | NR | .96 Y<br>.97 R<br>.94 I<br>(n=10) | NR | NR | .77-.89 OCR<br>CPRS-OC<br>NIMH-OC<br>.69-.77 NIMH-G | NR |
| LOI-CV (20) | | | | | | |
| Berg et al , 1988<br>(N= 5596) | .81 | NR | NA | NR | NR | NR |
| Flament et al , 1988<br>(N = 356) | NR | NR | NA | NR | 75% Sensitivity<br>84% Specificity<br>18% Predictive | NR |
| King et al , 1988<br>(N= 1602) | .76 | .71<br>(n=106) | NA | NR | NR | NR |
| | | | | | | |
| CY-BOCS | | | | | | |
| Hanna, 1995<br>(N= 31) | NR | NR | NR | NR | .93 NIMH-G | NR |

Note LOI-CV (44 and 20)=Leyton Obsessional Inventory- Child's Version (44 item and 20 item). NR=Not Reported; NA= Not Applicable; Y= Yes scores on LOI-CV, R=Resistance scores on LOI-CV, I=Interference scores on LOI-CV OCR=Obsessive Compulsive Rating (Rapoport et al., 1980); CPRS-OC=Comprehensive Psychopathological Rating Scale-Obsessive Compulsive Subscale (Thoren et al , 1980); NIMH-OC= NIMH-Obsessive-Compulsive Scale (Rapoport et al, unpublished); NIMH-G=NIMH-Global scale (Murphy et al., 1982) CY-BOCS = Child's Yale-Brown Obsessive Compulsive Scale

(c) the card sort method affords clinical utility through observations. The disadvantages are: (a) it does not distinguish between obsessions and compulsions; (b) it is time consuming to administer; and (c) there is potential for an overabundance of false positives and false negatives (Berg, 1989).

The Children's Yale-Brown Obsessive Compulsive Scale (CYBOCS). Like the LOI-CV, the CY-BOCS (Goodman, Rasmussen, Price, Riddle, & Rapoport, 1990) is a downward extension of the adult version, Y-BOCS. The CY-BOCS is a 19-item, clinician-administered scale, and is the most widely used outcome measure for childhood onset OCD. The first 10 items assess core OCD symptomatology (Items 1-5: obsessions; Items 6-10: compulsions), and comprise the total score. The remaining items assess associated features of OCD (5 items), global severity and improvement (2 items), reliability (1 item), and insight (1 item). All items are rated on a 5-point, Likert-type scale.

As is true for the LOI-CV (both 44-item and 20-item versions), there has been a paucity of research that has systematically investigated the technical properties of the CY-BOCS. This fact is surprising given its widespread use in drug outcome studies (Goodman & Price, 1992). A literature search was conducted through the Obsessive Compulsive Information center at the Dean Foundation located in Madison, Wisconsin by using the general descriptor, "Child Y-BOCS." This search revealed 24 articles. Only two of these 24 articles pertained to technical test information. Riddle et al. (1993) used 19 taped interviews and two blind raters to calculate inter-rater reliability, intraclass correlation coefficients, and a one-way ANOVA. Unfortunately, these statistics were not reported in this abstract, and it is unclear if a paper with these results was ever published. These authors indicated that validity measures would be computed in the future. Hanna (1995) reported validity measures of the CY-BOCS for a sample of 31 clinically referred children and adolescents with OCD. A Pearson product moment correlation of .93 was found for the CY-BOCS and the NIMH-G. Given that the technical properties of the CY-BOCS are relatively untested, it is difficult to make a judgment about the reliability, validity, and utility of this instrument. (see Table 1 for technical information on CY-BOCS.)

Despite the scarcity of technical information, the advantage of the CY-BOCS is its ability to assess several dimensions of OCD symptoms such as severity, interference, distress, degree of control, frequency, insight, avoidance, and improvement. The greatest concern, however, is that the authors did not retain the symptom list from the adult version; therefore, a clinician might inadvertently miss a key OCD symptom. Stallings and March (1995) conclude that there may be no clear advantage of the CY-BOCS over the Y-BOCS.

The LOI-CV and the CY-BOCS are probably the most common rating scales used for the assessment of OCD. These scales measure various characteristics of OCD and might be useful in certain situations. However, they share several limitations such as lack of adequate standardization as well as limited reliability and validity (specifically content validity and accuracy). Moreover, the limited and untested support for these devices attests to the importance of additional empirical work to extend the existing literature base and for alternative scale development.

## Direct Observation Assessment of OCD

Although there are no systematic direct observation measures currently used on a regular basis, this assessment method should be considered in the assessment of OCD. Direct observation strategies afford the observer the framework necessary to record behavior as it naturally occurs. There are several different methods for obtaining direct observations (i.e., frequency, duration, event, latency, and narrative). The chosen method will depend on the purpose of assessment. For example, if an observer is interested in the number of times the target child raises her/his hand, an event recording will likely be used given the discrete nature of hand raising. Direct observations are useful for the functional analysis of behavior in terms of antecedent, consequent, and sequential events. This method is the most direct method of behavioral assessment (Cone, 1978) and has the distinct advantage of having ecological validity. Currently, there are no systematic direct observation measures for diagnosis, treatment planning, or treatment monitoring of OCD. Work in this area is noticeably absent given researchers' pleas for direct observational measures (Berg, 1989; Taylor, 1995).

## Summary

The review of the childhood-onset OCD assessment literature revealed several shortcomings of the currently used assessment devices. These weaknesses are: (a) the children's scales have not been designed specifically for children in that they have been downward extensions of the adult scales; (b) the scales have untested reliability and validity; and (c) they are primarily clinician-administered interviews or rating scales. Additionally, direct behavioral observations for assessment and intervention monitoring have not been used on a consistent or frequent basis (Taylor, 1995).

A review of psychopharmacological treatment studies and psychosocial interventions for children who have OCD underscores the inadequacies of the current measures to reliably and accurately assess behavior change. In terms of both medication and psychosocial intervention, clinicians and researchers assert that there has been a positive change; however, they often fail to provide consumers with the data on what behaviors have changed.

Over the past decade, the investigation of successful treatment options for children and adolescents with OCD has multiplied. Fortunately, exposure and response prevention (E/RP) behavior therapy; drug therapy, or a combination of these therapies have afforded many children and adolescents some symptom relief; and for a few, full recovery (Leonard, Swedo, March, & Rapoport, 1995). Unfortunately, the assessment methodology of OCD significantly trails that of current treatment options.

Given the lack of any direct behavioral observation measure to monitor psychopharmacological and/or behavioral interventions in children and adolescents, the OCD-Direct Observation System (OCD-DOS) was created. The reliability and accuracy of the OCD-DOS will be established through an alternative scale validation model called the accuracy-reliability paradigm (Racine, 1994).

## Section Three: User's Guide to
## The Obsessive-Compulsive Disorder Rating Scale (OCD-RS) and
## Direct Observation System (OCD-DOS)

### OCD-RS

The OCD-RS is designed to measure how often a student exhibits specific, observable OCD behaviors in the classroom, home, and/or community. The OCD-RS also assess how these specific behaviors interfere with a student's success in a particular environment. This interference dimension lends itself nicely to the DSM-IV criteria as well as gives the evaluator an idea which behavior needs the most intervention. For purposes of this manual, the rater will rate four OCD target behaviors. These target behaviors are reassurance seeking, writing rituals, repeating, and symmetry. See Section Four for the operational definitions of these behaviors. A complete copy of the OCD-RS can be found in Appendix A.

### Directions for the OCD-RS

1. Familiarize yourself with the operational definitions of each target behavior.

2. Complete the identifying information sheet in as much detail as possible. Include any specific information regarding child, setting, and environmental characteristics.

3. Strategies for rating assessment include: (a) keeping a mental or written tally of the number of times a behavior occurred during the specified time

and then applying that number to the specific dimension of the rating scale: and (b) making a global judgment at the end of the specified time period.

4. At the end of the rating period, decide how often the student performed the behavior described.

- If student **never** engaged in this behavior, circle 0.
- If student performed this behavior **sometimes** (i.e., approximately 1-10 times/rating period), circle 1.
- If student performed this behavior **very often** (i.e., approximately 11+ times/rating period), circle 2.

5. In addition the frequency rating, rate how much each of these behaviors seemed to interfere with the student's classroom/home/community functioning.

- If the behavior did not seem to interfere with the classsroom/home/community performance and/or does not seem to cause the student distress when he/she is not allowed to perform the behavior, circle 0 (**No Interference**).
- If the behavior somewhat interfered with school/home/community performance and/or looks as though the student is in some distress when asked to stop the behavior, circle 1 (**Sometimes Interferes**).

- If the behavior interfered with school/home/community performance and/or student was in obvious distress if asked to stop behavior, circle 2 (**Interferes Most of the Time.**)

6. In addition to the target behaviors, record any additional OCD behaviors observed. You may start rating these during the current observation or choose to start rating these target behaviors for your next rating period. These additional observations are important to intervention monitoring given that OCD symptoms tend to wax and wane over time.

## OCD-DOS

The OCD-DOS is a structured direct behavior observation system for observing the **compulsive** behaviors of children and adolescents who have OCD. It should be noted that this observation system focuses only on the overt, compulsive behaviors commonly associated with OCD. Therefore, this observation system is designed to supplement other OCD assessment tools within a comprehensive assessment framework. It is intended to be used in a variety of settings such as school, home, and the community.

The OCD-DOS uses an event recording procedure to measure the frequency of target behaviors. The OCD-DOS requires the observer to record the target behavior with a tally mark for each occurrence of the behavior during the observation session. For purposes of this manual, the OCD-DOS focuses on the frequency of four target behaviors which are reassurance seeking, writing rituals, repeating, and symmetry. It should be noted that the rating scale and direct observation system are compatible in that they target the same behaviors. These operational definitions may be reviewed in Section Four. A complete copy of the observation form may be found in Appendix B.

## Directions for the OCD-DOS

1. Carefully familiarize yourself with the operational definitions of the four compulsive behaviors.

2. Complete the identifying information sheet in as much detail as possible. Include specific information regarding child, setting, and environmental characteristics. Also, record exact starting time.

3. For videotape observation, the observer may choose to observe for whole class period for a real time observation, or for a shortened amount of time (e.g., 20 or 30 minutes). Note time started and ended.

4. The OCD-DOS divides the observation period into five minute intervals. During the observation period, each target behavior is recorded in the appropriate time interval. For example, if Reassurance Seeking (RS) occurs at 10 minutes and 29 seconds, the observer would place a tally mark in the 10:01 to 15:00 interval column. If the observer observed two Writing rituals (W) during the 17th minute of observation, the observer would place two tally marks in the appropriate column. See the example below for an illustration.

| Min | 00-5:00 | 5:01- 10:00 | 10:01- 15:00 | 15:01- 20:00 | 20:01- 25:00 | 25:01- 30:00 | 30:01- 35:00 | 35:01- 40:00 | Total |
|---|---|---|---|---|---|---|---|---|---|
| RS | | | I | | | | | | |
| W | | | | II | | | | | |

5. The observer also should use the space provided at the bottom to record any additional observations (i.e., new compulsions observed, antecedents, consequents, sequential and/or environmental variables etc.).

6. The extra numbers on the OCD-DOS record form are for the observer to write in any other target behaviors.

7. Record the exact time at which the observation period ended.

Videotape Training

The videotape accompanying this manual provides a review of (1) OCD in children and adolescents, and (2) the accuracy-reliability paradigm. Additionally, it allows the assessor to calibrate his/her observation skills by reviewing the four operational definitions of OCD school behaviors highlighted in this manual and by observing two 40-minute classroom situations.

The videotape illustrates a case example of an adolescent girl who has problematic, overt compulsive behaviors in the school setting. Hypothetically, the diagnosis of OCD has already been rendered, and the purpose of the observations and ratings are for treatment monitoring. The operational definitions of these target behaviors are presented prior to the analog classroom situations (i.e., Language Arts Class #1 and #2).

## Section Four: Behavior Codes

Each behavior code category is operationally defined. In order to be an accurate observer or rater, observers/raters must memorize these definitions. Study the following definitions carefully and refer to them frequently during training. There is a self-administered quiz following the behavior codes. If you do not pass the quiz with 100% accuracy, please restudy the behavior codes and retake the quiz. The quiz answers may be found in Appendix C.

### Reassurance Seeking (RS)

This type of behavior includes excessive questioning of directions, assignments, activities, time limits, classroom rules, etc. Basically, this person's behavior conveys a need for reassurance by someone else. RS is distinguished from typical questions primarily by quantity. These RS questions are different from that of clarification questions in that the listener (i.e., observer/rater) has the intuitive sense of what the answer is.

### Writing Rituals (W)

This behavior consists of excessive writing, rewriting and/or retracing. Difficulty with writing could include the following behaviors: elaborate correction, writing/rewriting, tracing certain letters, and starting whole assignment over.

### Repeating (R)

This behavior is defined as having to do particular actions over and over again. Repeating behaviors include but are not limited to the following: getting school work/supplies out of and putting them back into backpack; having to come through the door several times in a certain way, blowing on fingers before one picks up a pencil or turns a page, standing up and sitting down, tapping fingers a certain number of times before starting work, touching a glass to one's lips a certain amount of times, clicking the pen a number of times, reading over and over. One episode of repeating is coded as one event recording. That is, if student engages in R by standing up and sitting down three times, that would be coded as one event.

**Symmetry (SY)**

This behavior conveys a need for exactness or order. Symmetry behaviors could include the following: lining up a book or paper with the edge of a desk, making sure a writing implement is placed exactly in the center of the desk, arranging and rearranging supplies. One episode of symmetry is coded as one event recording. That is, if student engages in SY by lining up her book with the edge of a desk several times before stopping, that would be coded as one event.

# Behavior Codes Quiz for the OCD DOS and the OCD RS

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Reassurance Seeking* (RS).

1.______ Student asks the teacher to clarify an assignment in further detail by asking micro-specific questions. Classmates become irritated with the student.

2.______ Student asks a peer for the assignment because she was too busy copying the last assignment down.

3.______ Student asks two questions about a particular assignment. Then, two more students ask similar questions for assignment clarification.

4.______ Student wants to know from the teacher if she is doing an assignment correctly. She gets visibly upset when teacher tries to move on.

5.______ Student asks another student if she is doing the assignment right, while all the rest of the students seem to be working on their assignments with no trouble or confusion.

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Writing Rituals* (W).

1.______ Student appears to be writing very slowly, but you can't really see what she is doing on the paper.

2.______ Student is erasing excessively, and then starts writing again.

3.______ Student rips paper out of her spiral notebook, and starts the assignment over again. She appears in distress.

4.______ Student appears to be tracing a letter over and over. You can actually tell that she is tracing by looking at her making the same movement over and over again.

5.______ Student appears to be tracing a letter over and over. You can't quite be sure that she actually is tracing and re-tracing.

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Repeating* (R).

1.______ Student runs her fingers through her hair several times while she is going to retrieve a book. She does not look as though she realizes she is doing this behavior.

2.______ Student bites her nails while she is reading a book in class.

3.______ Student is entering class, but appears to have forgotten something. This student retreats back through the doorway and tries to enter through the door. Now, the student has a grimace on her face, and tries to come through the doorway again. Finally, the student takes her seat.

4.______ Student has a drink in the lunch room. She taps the glass to her lips seven times before drinking from the glass.

5.______ Student has to stand up three times before being able to leave her classroom.

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Symmetry* (SY).

1.______ Student opens her book to the assigned page and proceeds to smooth down her pages.

2.______ Student gets her supplies out of her backpack and looks as if she has to place them in their correct places.

3.______ Student has her spiral notebook resting on her lap and her pen haphazardly placed on her desk along with her assignment notebook.

4.______ Student has nothing on her desk.

5.______ Student places her spiral notebook on her desk, and pays careful attention to lining up the side of her spiral with the side of her desk.

## Section Five: A Note to Research Participants

Experts

1. You have been mailed the videotape, rating and observation manual, and the OCD-RS and OCD-DOS forms.

2. Please read the manual carefully and familiarize yourself with the target behavior definitions and directions. While there exists some debate in the OCD field on which behaviors are actually OCD versus those which are not OCD, our instruments are not intended to render a diagnosis of OCD, but rather to monitor treatment of OCD. Therefore, when you watch the videotape please focus on the operational definitions of the target behavior rather than is it or is it not OCD.

3. Next, watch the first portion (20 minutes) of the videotape. When you get to the analog classroom scenarios, stop your VCR. Now, you need to have two OCD-DOS and OCD-RS because you will need one of each for each classroom scenario on the videotape (i.e., Language Arts Class #1 and #2).

5. For the first classroom scenario (i.e., Language Arts Class #1), please observe using the OCD-DOS starting at 20:00 and stop at 45:00. These intervals are labeled at the top of your OCD-DOS.

6. You can complete the OCD-RS after you finish the OCD-DOS.

7. Please do not stop and rewind the VCR, it is our hopes that your observation will simulate as much as possible a naturalistic observation.

8. When you have completed the first observation and rating, you can observe and rate the next classroom situation (Language Arts Class #2) starting at 1:23 and ending at 26:23. These intervals are labeled at the top of the OCD-DOS. Again please do not stop and rewind while you are completing your observation form.

9. Once you have completed your two observations and ratings, please return the materials and the videotape to Caroline N. Racine in the mailer provided.

We certainly appreciate your time and effort. Hopefully, your contribution will result in better assessment methods for children and adolescents with OCD. When we receive all your materials, a check of $500 will be forthcoming under separate cover. Thanks.

Graduate Student Observers

1. You will receive an observation training session that will last approximately 3 hours. You will achieve a mastery criterion of 85% agreement using the OCD-DOS and other event recording observation systems.

2. During the training period, you will be required to read the Rating and Observation Manual and familiarize yourself with the target behaviors.

3. Once this initial training is completed, you will be randomly assigned to one of two groups. What group you are in will dictate what task you do for that week.

4. The data collection will take 6 weeks and last for 30 minutes each time we meet.

5. We will meet for the next 6 Mondays at 4:15 PM in 319.

## Section Six: Technical Information

### Reliability

To be determined

### Accuracy

The most glaring shortcoming of all current OCD assessment devices has been the failure to use appropriate validity indices. This lack of appropriate validity estimates for OCD assessment devices significantly hinders the utility of these instruments. All of the OCD measures that report validity estimates have used a very traditional method for establishing content, construct, and criterion-related validity. In essence, investigators have tried to validate OCD assessment instruments by using a traditional scale validation paradigm that seeks to assess the existence of a hypothetical construct, or latent trait. While this traditional validation paradigm represents one way to establish the validity of instruments, perhaps an alternative scale validation model may be more useful in establishing the validity of OCD assessment measures. In this context, an effort should be made to validate the OCD assessment devices as measures of behavior instead of measuring a latent trait, or state of being. In other words, OCD can (and should) be considered a behavioral response class consisting of numerous individual, molecular behaviors (i.e., obsessions and compulsions). The purpose of validating OCD assessment measures as behavioral assessment measures is to ensure that the resulting scores describe actual, observable behavior (i.e., motor responses, cognitive and affective events, and physiological responses).

Foster and Cone (1995) assert that two types of representational validity, content validity and accuracy, should be established for behavioral measures. Content validity assesses the extent to which the instrument measures what it purports to measure. Specifically, content validity examines the extent to which a measure operationally defines the behavior/response class that is being assessed. Also, scores obtained from the measure should depict the actual behavior. Accuracy is a measure of the instrument's ability to reflect "true" behavior. Establishing an instrument's accuracy is done by comparing scores on a given instrument to those of an incontrovertible index. An incontrovertible index is an illustration of the target behavior(s) that the investigator deems as "truth." This incontrovertible index can be generated by mechanically generated responses, naturalistic observations, and/or controlled stimuli (i.e., videotaped naturalistic behavior or scripted performance). Once accuracy of the instrument is established, convergent and discriminant validity are no longer necessary (Foster & Cone, 1995).

Given that the psychometric properties of the various childhood-onset OCD scales are suspect, there is a need to develop a scale that is sensitive to the topography of a certain behavior. Accuracy is a concept that is closely related to reliability and validity, yet has rarely been addressed in the evaluation of rating scales or direct observation measures. Cone (1981 ) indicated that the establishment of an instrument's accuracy is the most important feature. Unfortunately, researchers have often assumed accuracy from inter-rater reliability or test-retest reliability indices. However, a measure that is reliable is not necessarily accurate, but a measure that is accurate is inevitably reliable. Cone (1992) asserts, "To establish an instrument's accuracy, whether for occurrence, cross-setting, or other, it is necessary to have two things: (a) a set of rules/procedures, preferably written, for using the instrument, and (b) an incontrovertible index against which to compare data produced by the instrument" (p. 24). Thus, an incontrovertible index will be established for purposes of this research by using videotaped vignettes of naturalistic behavior and/or scripted performances from children who exhibit the various OCD subtypes. Then, experts in the field of OCD will view these videotapes and complete the newly created rating scale and direct observation measure. If the experts converge on what they are observing, then the scale will be accurate. Additionally, written, explicit instructions will be developed in accordance with Cone's recommendations. Instructions will be an important part because they are rarely used with existing instruments (Taylor, 1995).

Other accuracy indices will be determined from current research project.

## Behavior Codes Quiz for the OCD DOS and the OCD RS

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Reassurance Seeking* (RS)

1.______ Student asks the teacher to clarify an assignment in further detail by asking micro-specific questions. Classmates become irritated with the student.

2.______ Student asks a peer for the assignment because she was too busy copying the last assignment down.

3.______ Student asks two questions about a particular assignment. Then, two more students ask similar questions for assignment clarification.

4.______ Student wants to know from the teacher if she is doing an assignment correctly. She gets visibly upset when teacher tries to move on.

5.______ Student asks another student if she is doing the assignment right, while all the rest of the students seem to be working on their assignments with no trouble or confusion.

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Writing Rituals* (W).

1.______ Student appears to be writing very slowly, but you can't really see what she is doing on the paper.

2.______ Student is erasing excessively, and then starts writing again.

3.______ Student rips paper out of her spiral notebook, and starts the assignment over again. She appears in distress.

4.______ Student appears to be tracing a letter over and over. You can actually tell that she is tracing by looking at her making the same movement over and over again.

5.______ Student appears to be tracing a letter over and over. You can't quite be sure that she actually is tracing and re-tracing.

BEST COPY AVAILABLE

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Repeating* (R).

1. ______ Student runs her fingers through her hair several times while she is going to retrieve a book. She does not look as though she realizes she is doing this behavior.

2. ______ Student bites her nails while she is reading a book in class.

3. ______ Student is entering class, but appears to have forgotten something. This student retreats back through the doorway and tries to enter through the door. Now, the student has a grimace on her face, and tries to come through the doorway again. Finally, the student takes her seat.

4. ______ Student has a drink in the lunch room. She taps the glass to her lips seven times before drinking from the glass.

5. ______ Student has to stand up three times before being able to leave her classroom.

Instructions

For each behavioral description below, decide if the behavior is an example (E) or non-example (N) of *Symmetry* (SY).

1. ______ Student opens her book to the assigned page and proceeds to smooth down her pages.

2. ______ Student gets her supplies out of her backpack and looks as if she has to place them in their correct places.

3. ______ Student has her spiral notebook resting on her lap and her pen haphazardly placed on her desk along with her assignment notebook.

4. ______ Student has nothing on her desk.

5. ______ Student places her spiral notebook on her desk, and pays careful attention to lining up the side of her spiral with the side of her desk.

# Appendix D

## Raw Data

Raw Data
Videotape #1

| | OCD-RS | | | | OCD-DOS | | | |
|---|---|---|---|---|---|---|---|---|
| Participant | RS | W | R | SY | RS | W | R | SY |
| 1 | 2 | 2 | 1 | 1 | 22 | 18 | 2 | 4 |
| 2 | 2 | 2 | 1 | 1 | 26 | 17 | 4 | 6 |
| 3 | 2 | 2 | 1 | 1 | 21 | 18 | 3 | 6 |
| 4 | 2 | 2 | 1 | 1 | 22 | 14 | 3 | 3 |
| 5 | 2 | 1 | 1 | 0 | 20 | 8 | 2 | 2 |
| 6 | 2 | 2 | 1 | 1 | 19 | 11 | 2 | 2 |
| PS | 2 | 2 | 1 | 1 | 15 | 16 | 3 | 4 |
| E1 | 2 | 2 | 1 | 2 | 23 | 14 | 6 | 11 |
| E2 | 1 | 1 | 1 | 1 | 20 | 16 | 7 | 3 |
| E3 | 2 | 2 | 1 | 1 | 20 | 12 | 6 | 6 |
| E4 | 2 | 2 | 1 | 1 | 14 | 17 | 3 | 2 |
| E5 | 2 | 2 | 1 | 1 | 13 | 16 | 3 | 2 |
| E6 | 2 | 2 | 1 | 1 | 25 | 15 | 9 | 5 |
| E7 | 2 | 2 | 1 | 1 | 20 | 15 | 6 | 5 |
| E8 | 2 | 2 | 1 | 1 | 21 | 17 | 3 | 9 |
| E9 | 1 | 2 | 2 | 1 | 7 | 11 | 12 | 7 |

Videotape #2

| | OCD-RS | | | | OCD-DOS | | | |
|---|---|---|---|---|---|---|---|---|
| Participant | RS | W | R | SY | RS | W | R | SY |
| 1 | 1 | 1 | 0 | 1 | 10 | 3 | 0 | 8 |
| 2 | 2 | 2 | 0 | 1 | 16 | 4 | 0 | 8 |
| 3 | 2 | 1 | 0 | 2 | 12 | 3 | 0 | 10 |
| 4 | 2 | 1 | 0 | 2 | 18 | 1 | 0 | 11 |
| 5 | 2 | 1 | 0 | 1 | 13 | 0 | 0 | 10 |
| 6 | 2 | 1 | 0 | 2 | 16 | 3 | 0 | 11 |
| PS | 1 | 1 | 0 | 1 | 9 | 3 | 0 | 7 |
| E1 | 1 | 1 | 0 | 1 | 10 | 3 | 0 | 10 |
| E2 | 1 | 1 | 1 | 1 | 12 | 3 | 5 | 7 |
| E3 | 2 | 1 | 0 | 2 | 13 | 3 | 0 | 7 |
| E4 | 1 | 1 | 0 | 1 | 3 | 9 | 0 | 3 |
| E5 | 0 | 1 | 1 | 1 | 0 | 4 | 0 | 9 |
| E6 | 2 | 2 | 0 | 1 | 10 | 11 | 7 | 8 |
| E7 | 2 | 1 | 1 | 1 | 14 | 4 | 6 | 8 |
| E8 | 2 | 1 | 0 | 1 | 15 | 5 | 0 | 6 |
| E9 | 1 | 1 | 1 | 1 | 8 | 5 | 3 | 9 |